

RESEARCH ARTICLE

Clutter negotiating ability in an ensemble of forest interior bats is driven by body mass

Juliana Senawi^{1,2,3,*} and Tigga Kingston³**ABSTRACT**

Differences in wing morphology are predicted to reflect differences in bat foraging strategies. Experimental tests of this prediction typically assess the relationship between wing morphology and a measure of flight performance on an obstacle course. However, studies have lacked measures of obstacle avoidance ability true scores, which may confound interpretation of ability across the range of presented tasks. Here, we used Rasch analysis of performance in a collision-avoidance experiment to estimate the ability of bat species to fly through vegetative clutter. We refer to this latent trait as ‘clutter negotiating ability’ and determined the relationships between clutter negotiating ability and wing morphology in 15 forest insectivorous bat species that forage in the densely cluttered rainforests of Malaysia. The clutter negotiating ability scores were quantified based on individual responses of each species to 11 different obstacle arrangements (four banks of vertical strings 10–60 cm apart). The tasks employed for the collision-avoidance experiment were reliable and valid, although Rasch analysis suggested that the experiment was too easy to discriminate completely among the 15 species. We found significant negative correlations between clutter negotiating ability and body mass, wingspan, wing loading and wing area but a positive significant correlation with wingtip area ratio. However, in stepwise multiple regression analyses, only body mass and wing loading were significant predictors of clutter negotiating ability. Species fell into clusters of different clutter negotiating ability, suggesting a potential mechanism for resource partitioning within the forest interior insectivorous ensemble.

KEY WORDS: Item response theory, Maneuverability, Flight, Rasch analysis, Resource partitioning, Wing morphology

INTRODUCTION

In insectivorous bat assemblages, partitioning of food resources is effected largely by physical and sensory access to the insect prey base. Differences in wing morphology influence access to particular foraging habitats and the prey capture strategy deployed within them (Norberg and Rayner, 1987; Fenton, 1990; Norberg, 1994, 1998). The functional relationship between foraging habitat and wing morphology in bats is conventionally expressed by differences in wing aspect ratio ($\text{wingspan}^2/\text{wing area}$), wing loading [$(\text{mass} \times \text{gravity})/\text{wing area}$] and wingtip shape (Norberg and

Rayner, 1987). These parameters influence flight speed and maneuverability (the space required by a bat to alter the flight path) (Aldridge, 1987; Norberg and Rayner, 1987; Aldridge and Brigham, 1988). In insectivorous bat assemblages, flight morphology is commonly used to infer three main foraging ensembles – uncluttered, semi-cluttered and cluttered – based primarily on how wing parameters influence the ability of bats to fly and forage in cluttered environments (Aldridge and Rautenbach, 1987; Fenton, 1990; McKenzie et al., 1995; Schnitzler and Kalko, 1998). Bats that forage in uncluttered environments (‘open-space’ bats; *sensu* Schnitzler and Kalko, 1998), e.g. above the forest canopy or in open habitats, typically combine high wing loading to facilitate fast flight with high aspect ratios (long, narrow wings) (Norberg and Rayner, 1987; Norberg, 1990; Swartz et al., 2012). Species that forage in semi-cluttered environments, such as tree fall gaps and forest edges (‘edge/gap’ bats), typically combine high wing loading and low aspect ratio or low wing loading and high aspect ratio (Norberg and Rayner, 1987) to achieve greater maneuverability, although flight speeds are lower. ‘Narrow-space bats’ forage in highly cluttered environments, such as the interior of a rainforest. Species within this ensemble maximize maneuverability with low aspect ratio and low wing loading, and flight is slow.


Although there are clear differences in wing morphology among ensembles, the extent and functional significance of morphological variability within ensembles is poorly known. Yet, differences in performance within an ensemble could partition foraging space and facilitate species coexistence. This is of particular interest in the Palaetropics, where the narrow-space ensemble of insectivorous bats that forage in the cluttered interior of rainforests can include more than 30 species of five families (Kingston et al. 2003, 2006). Although prior studies have demonstrated that coexistence in this forest interior ensemble might be facilitated through differences in acoustic prey detection (Kingston et al., 2000; Schmieder et al., 2012), and the size and hardness of prey that can be manipulated (Senawi et al., 2015), flight performance, particularly the ability to fly through clutter, could provide another dimension by which species partition resources.

Studies have experimentally tested predictions about the relationship between bat morphology and flight performance in clutter using obstacle courses (e.g. Aldridge, 1986; Aldridge and Rautenbach, 1987; Jones et al., 1993; Stockwell, 2001). The courses require bats to fly through sets of vertical rods or strings, and the degree of clutter is manipulated by increasing or decreasing the distance between the rods/strings. Performance is then scored based on collision avoidance and correlated with aspects of flight morphology. The performance is often described in terms of maneuverability, which was originally defined by Norberg and Rayner (1987) as one of two measures of turning performance, or the ability to change flight direction, and referred to the minimum space required for a turn at a given speed. The other measure of turning performance is agility, which describes the rate at which a

¹School of Environmental and Natural Resource Sciences, Faculty of Science and Technology, Universiti Kebangsaan Malaysia, 43600 Bangi, Selangor, Malaysia.

²Institute for Environment and Development (LESTARI), Universiti Kebangsaan Malaysia, 43600 Bangi, Selangor, Malaysia. ³Department of Biological Sciences, Texas Tech University, Lubbock, TX 79409, USA.

*Author for correspondence (julianas@ukm.edu.my)

 J.S., 0000-0002-5689-5718; T.K., 0000-0003-3552-5352

turn can be initiated. A high score on an obstacle course could potentially be achieved through high maneuverability or agility, or a combination of the two. Moreover, the flight behaviors associated with these measures, and with the ability to pass through gaps between obstacles that measure less than wingspan, involve complex wing and body kinematics (Bergou et al., 2015). Navigating through clutter may thus integrate many aspects of flight achieved through diverse kinematic mechanisms. However, from an ecological and functional perspective, it is the overall ability to fly through clutter without colliding with obstacles that is of interest. This ability, which we call ‘clutter negotiating ability’, can be conceptualized as a latent construct or variable – an unobservable characteristic or trait – that can be evaluated through latent trait models. Latent trait models aim to measure the underlying ability (or trait), in this case clutter negotiating ability, that is producing performance in a test (in this case the collision-avoidance experiments), rather than measuring the performance per se.

Rasch analysis is a latent trait modelling approach widely used to develop scales that measure performance, attitudes and perceptions in human sciences (see Bond and Fox, 2015; Boone et al., 2014; Boone, 2016; Engelhard, 2012). It provides a mechanism for estimating the reliability and validity of observed scores of a performance in a test and determining the true score (latent ability score) of subjects in experiments comprising multiple tasks. In collision-avoidance experiments, the distance between the rods/strings represents a task and collectively the tasks make up an experiment that describes the performance of the bats. To date, performance scores have been quantified directly from observed scores of tasks from the experiment. This may compromise interpretations of performance because true score theory states that the observed score is equal to the sum of the true score or true underlying ability plus the measurement error associated with estimating the observed score, or:

$$X = T + e_x, \quad (1)$$

where X is an observed score, T is a true ability and e_x is measurement error (Allen and Yen, 2002). A score that has no measurement error (i.e. true score) is perfectly reliable whereas a measure that has no true score (i.e. all variability is attributable to measurement error) has zero reliability. Thus, reliability describes the consistency or reproducibility of the test score (Trochim, 2006; DeVon et al., 2007), which indicates the quality of the test. Therefore, the observed scores of a flight performance experiment are not synonymous with the ability score because observed scores will always depend on the selection of flight performance tasks that comprise the experiment over which their ability score is defined, i.e. the number, range and values of the inter-string distances. An experiment of five tasks with inter-string distances increasing from 5 to 25 cm (5 cm intervals) is not the same as one of five tasks with inter-string distances increasing from 10 to 50 cm. Conceptually, observed scores are test dependent but ability is test independent (Lord, 1953; Hambleton and Jones, 1993). In practice, bats are more likely to have lower true scores on difficult tasks and higher true scores on easier tasks, but their ability remains constant over any task that might be built to measure ability (Wright and Stone, 1979). Ability scores or estimated ability scores that are independent of the particular flight performance tasks would be of value because they would allow an unbiased comparison among species. Validity provides another indication of test quality, referring to whether or not the experiment (or individual tasks) measures what it claims to measure (Messick, 1995; Borsboom et al., 2004).

Rasch analysis is built around a statistical logistic response model and specifies that each task response is taken as an outcome of the linear probabilistic interaction of a subject’s ability and a task’s difficulty such that the scale is a unidimensional measure of ability (Wright and Masters, 1982; Andrich, 1988). Unidimensionality of a scale indicates that tasks used for the scale reflect a single construct of function. Under this construct, tasks vary from easy to difficult, and test subjects can be ranked based on ability from less to more able on the same scale. For example, the scale of the tasks in a collision-avoidance experiment should be able to assess clutter negotiating ability of each insectivorous bat species within a specified environment. Easy tasks can be performed by bats with almost all clutter negotiating ability levels, whereas only more able bats are likely to successfully perform the difficult tasks. A set of tasks that focuses too heavily on the most difficult tasks, representing the most cluttered environments, may not be a good measure of a bat’s ability to express their clutter negotiating ability because the task may overlook the bat’s ability in the least cluttered environments and vice versa.

To understand the relationships among clutter negotiating ability, wing morphology and foraging ecology within a foraging ensemble, we tested the clutter negotiating ability of 15 insectivorous bat species from Malaysia with a collision-avoidance experiment. Our collision-avoidance experiment essentially focused on the ecological challenge of passing through clutter without impacts (touches and breaks of strings) that, if repeated in the wild, would cause wing damage and ultimately constrain microhabitat use, leading to resource partitioning of the forest interior ensemble through differences in foraging space. We used Rasch analysis to generate true clutter negotiating ability scores and assess the reliability and validity of the experiment. Bat species were of three families but belonged to the same ensemble: forest interior insectivores that forage in the cluttered environments of the rainforest understory. Here, three specific questions were addressed. (1) Are collision-avoidance experiments reliable and a valid test of clutter negotiating ability? (2) If so, what is the relationship between clutter negotiating ability and wing morphology within the forest interior foraging ensemble? (3) Do differences in wing morphology and clutter negotiating ability provide a potential mechanism for resource partitioning within the forest interior ensemble?

MATERIALS AND METHODS

Study site and species

The study was conducted in Krau Wildlife Reserve, an area of 62,395 ha of continuous old-growth forest, located in the state of Pahang, Malaysia (DWNP-DANCED, 2001). The reserve supports at least 55 insectivorous bat species of eight families with about 30 of these species foraging in the forest interior (cluttered environment) (Kingston et al., 2006). However, this study focused on the most common insectivorous bats species caught in the forest interior habitat in a 300 ha area around Kuala Lompat Research Station (KLRS: 3°43’N, 102°17’E) on the eastern edge of the reserve. The elevation at KLRS is approximately 50 m a.s.l. with vegetation classified as lowland evergreen mixed dipterocarp forest (Hodgkinson et al., 2004). Insectivorous bats were captured in the forest understory using four-bank harp traps (Francis, 1989) positioned across trails. Harp traps were attended at 21:00 h and captured bats were held individually in cloth bags and identified from morphological characters following Kingston et al. (2006). Only adult, non-pregnant individuals were used in this study. All captured individuals were banded and released at the point of

capture within 12 h after all essential measurements and experiments were completed. Trapping and flight performance experiments were conducted between May and July 2010 and July and September 2011. All procedures were approved by Department of Wildlife and National Park Peninsular Malaysia and Texas Tech University, Lubbock (IACUC 10014-04).

Morphological measurements

For each captured individual we measured the length of the forearm from the outside of the elbow to the outside of the wrist in the bent wing with a dial caliper to the nearest 0.1 mm. We also measured the body mass (M) using Pesola scales (Pesola AG, Baar, Switzerland) to the nearest 0.25 g. Photographs of the wings with both wings fully extended were taken for each individual captured, using a Canon PowerShot G10 digital camera (Canon USA Inc., Melville, NY, USA) mounted on a quadra-pod (Forensic Imaging Inc., Victor, NY, USA). The bat was placed on a drafting mat and adhesive packaging tape was applied to restrain its extended wings and wing tail membrane while the photograph was taken (see McKenzie et al., 1995, Saunders and Barclay, 1992). This procedure took less than 1 min for each individual. Using image processing and analysis in Java (ImageJ version 1.48, National Institutes of Health, Bethesda, MD, USA), wing length and wing area (Fig. 1) were measured for each individual. From these measurements, we calculated the following additional wing parameters for each bat (after Norberg and Rayner, 1987). Wing span, B :

$$B = 2 \times \frac{1}{2}B, \quad (2)$$

where wing span is the distance between wingtips when the wings are fully extended; $\frac{1}{2}B$ is half the wingspan of the bat determined from wing photographs (Fig. 1). Wing area, S :

$$S = 2 \times \frac{1}{2}S, \quad (3)$$

where wing area is the area of both wings and the body surface area, excluding the head, when the bat's wings are fully extended; $\frac{1}{2}S$ is half the wing area of the bat determined from wing photographs

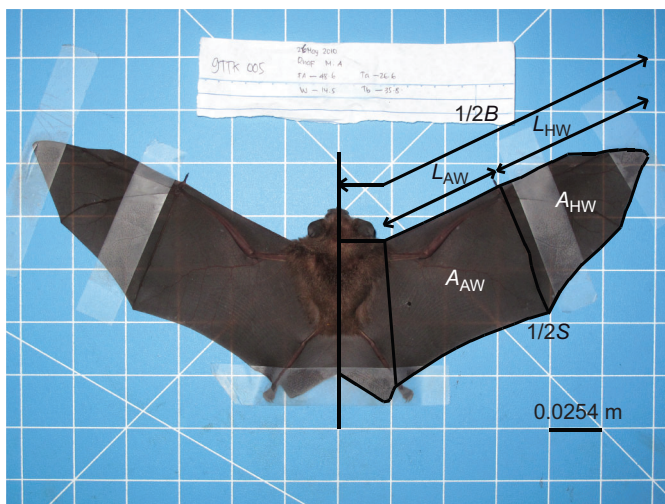


Fig. 1. Wing measurements. Photograph of a bat (*Rhinolophus affinis*), indicating the wing dimensions measured using ImageJ. $\frac{1}{2}B$, half wingspan; L_{AW} , arm-wing length; L_{HW} , hand-wing length; A_{AW} , arm-wing area; A_{HW} , hand-wing area; and $\frac{1}{2}S$, half wing area ($A_{AW}+A_{HW}$ +area between the midline of the body and the proximal edge of the A_{AW}).

(Fig. 1). Aspect ratio, AR:

$$AR = \frac{B^2}{S}, \quad (4)$$

where B is wing span and S is wing area. Wing loading, WL:

$$WL = \frac{Mg}{S}, \quad (5)$$

where M is body mass, S is wing area and g is acceleration due to gravity, estimated at 9.8 m s^{-2} . Wingtip length ratio, TL:

$$TL = \frac{L_{HW}}{L_{AW}}, \quad (6)$$

where L_{HW} is hand-wing length and L_{AW} is arm-wing length, both of which were determined from wing photographs (Fig. 1). Hand-wing length was measured from the proximal end of the carpus to the distal end of the third finger. Arm-wing length was measured from proximal end of the propatagium to the distal end of the carpus in fully extended wings. Wingtip area ratio, TS:

$$TS = \frac{A_{HW}}{A_{AW}}, \quad (7)$$

where A_{HW} is hand-wing area and A_{AW} is arm-wing area, both of which were determined from wing photographs (Fig. 1). Hand-wing area is the surface area of the hand-wing measured between the midline of the fifth finger and the wing-tip. Arm-wing area is the surface area of the arm wing measured between the proximal end of the propatagium and the proximal end of the wing membrane near the foot and midline of the fifth finger (Fig. 1). Wingtip shape index, I :

$$I = \frac{TS}{TL - TS}, \quad (8)$$

where TS is a wingtip area ratio and TL is wingtip length ratio.

Experimental design

Collision-avoidance experiments were conducted in a flight cage located at the forest edge. The cage was 3 m wide \times 3 m tall \times 10 m long and the walls made of $\frac{1}{2}$ inch steel hexagonal twist chicken wire. Experiments were conducted between 22:00 h and 06:00 h. Flight performance was evaluated using an obstacle course of four banks of vertical strings that were suspended in the center of the flight cage (Fig. 2). The distance between strings could be set at 10, 15, 20, 25, 30, 35, 40, 45, 50, 55 or 60 cm and each bank was set to the same inter-string distance for a given test but off-set one bank to another (Fig. 2). Resetting the distances took several hours, so a different inter-string distance (D) was tested each night, with inter-string distance progressing from 10 to 60 cm, then back to 10 cm. This rotation was continued for 16 sets (176 nights). These distances were considered representative of the range of clutter encountered by the insectivorous bat species tested when foraging in the forest understory around KLRS.

The banks were made from curtain rails to make it easier to change the inter-string distance. Each vertical string consisted of a mason line twisted nylon string (diameter 1.5 mm), secured to the top and bottom rails by curtain hooks. The line was cut between 5 and 20 cm from the bottom rail and the descending and ascending pieces of line were attached to a magnet and thumbtack, respectively, to provide a magnetic attraction that held the string in place vertically. Hitting the string hard enough to break the magnetic adhesion might reflect physical consequences for the bat in the wild. In this experiment, strong contact with the string broke the attraction, leaving the long descending piece of string hanging, while

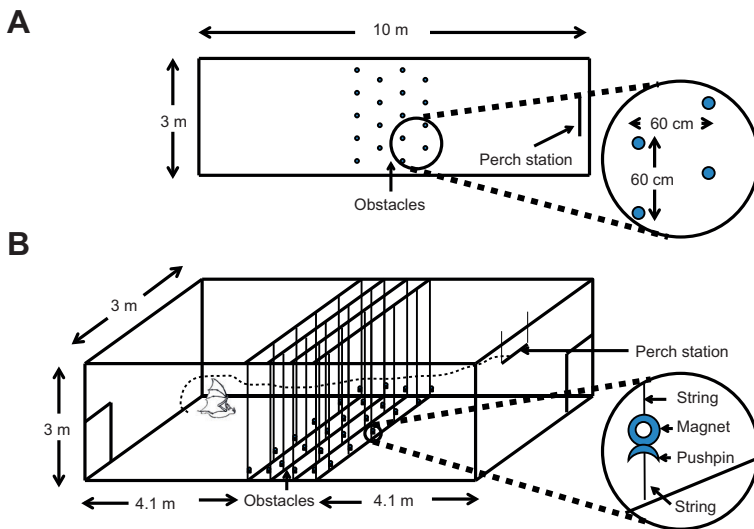


Fig. 2. Experimental setup. Each bat was presented with four rows of obstacles with evenly spaced inter-string distance but off-set from one row to another. The bat began the trial on a perch station and negotiated the obstacle course 10 times to complete one test (10 passes per test). (A) Top view. (B) Side view.

the small ascending piece dropped to the ground. When the contact was not strong enough, touches were indicated by a movement of the string while remaining attached via the magnetic attraction. This mechanism gave an easy indication that a bat had hit the string as fast-flying bats were expected and the experiments were conducted in natural darkness. Captured bats were presented with an inter-string distance set for that night and had to negotiate the obstacle course for 10 passes (counted as one test). At the start of the test, the bat was left undisturbed, hanging at the perch station (Fig. 2), until it flew through the obstacle course. A bat's observed score was quantified after each pass by observation of obstacles touched and manual counting of broken magnetic connections (which were reset between passes). Two observers were present for this experiment, one designated for recording the score and another one restoring the magnetic connections as needed. Both observers used a headlamp equipped with a red filter to avoid disturbing the bats, which were free to continue with the pass attempts, or rest between them.

From an ecological perspective, clearing is substantially better than touching because in the wild repeated touching is likely to cause cumulative and unsustainable wing damage, and ultimately constrain microhabitat use. So our scoring system heavily emphasized clearing the strings of a bank, for which the bat was awarded 5 points, whereas touching only scored 1 point and break 0. Thus, across all four banks, potential scores ranged from 0 to 20. In polytomous Rasch models, responses for the task must be scored with successive integers starting at zero (Andersen, 1977; Andrich, 1978; Salzberger, 2010); in essence, scores are viewed as a count of all thresholds that a respondent (bat species) has passed, provided the scoring adequately reflects the order of the thresholds (Andrich, 1995a,b). Therefore, we converted the total scores derived from passing through the four banks to a scale of 0 to 15 where 0=refusal and 15=cleared all banks without touching or breaking any strings. Scores in between depended on the combination of breaks, touches and clear runs across the four banks (Table S1). The observed score between 0 and 15 was then used in subsequent analysis.

Rasch analysis – modeling clutter negotiating ability score from observed collision-avoidance score

To convert the observed scores from individual tests at set inter-string distances to species clutter negotiating ability scores, we first assumed that individuals of the same species had similar clutter negotiating ability. Then, we determined the mode of the observed

score for tests of each individual (10 passes per test). The mode was chosen as a measure of the central tendency to avoid the potential influence of low scores from passes at the beginning of the test when the bats were familiarizing themselves with the course, or at the end if they began to tire. If multiple individuals of the same species were tested at the same inter-string distance, we took the mode of the observed scores across individuals. The output going into the Rasch analysis was therefore a single observed score for each species at each inter-string distance.

Rasch analysis was used to convert the observed inter-string distance scores to overall clutter negotiating ability scores for each species and to score the inter-string distance for difficulty. First, a scale of the inter-string distance was constructed by comparing the response patterns of species from the entire sample, creating a clutter negotiating ability scale for the collision-avoidance experiment conducted. In this study, the scale is a linear transformation of the Rasch logit scale to fit a -10 to 10 scale (we standardized the measures to have zero mean and standard deviation of one with $USCALE=0.3170$ and $UMEAN=-0.2894$). The logit is the natural logarithm of the odds of a bat being able to negotiate a particular inter-string distance. Then, species were placed on the scale according to their clutter negotiating ability score (true score) derived from all tasks in the experiment. The length of the clutter negotiating ability scale determines the range of clutter negotiating ability for all the species tested and the range of task difficulties was expected to be in order of inter-string distance, with the easiest task ($D_{60\text{cm}}$) at the base of the scale and the hardest task ($D_{10\text{cm}}$) at the top of the scale. Disordering would indicate that the clutter negotiating scale does not work as intended for the sample of species tested.

Finally, the quality of the clutter negotiating ability scale was determined with goodness-of-fit statistics. Specifically, we examined whether the clutter negotiating scale constructed met the criteria for unidimensionality and a hierarchical structure of inter-string distance difficulty. The assessment of goodness-of-fit statistics was based on the mean-square (MNSQ) value (size of randomness, i.e. the degree of distortion of the measurement system) of outfit and infit statistics. Outfit or outlier-sensitive fit statistics are more sensitive to unexpected scores of species for inter-string distances that are relatively very easy or very hard for them. Infit or inlier pattern-sensitive fit statistics are more sensitive to unexpected patterns of species for the inter-string distances. The expected pattern is that in which the order of scores follows the order

of task difficulty (i.e. higher score on easier tasks). $MNSQ > 1.5$ indicates possible misfits between a specific item (bat clutter negotiating ability score or inter-string distance difficulty score) and the rest of items in the scale (Linacre, 2019). To reduce the possibility of misinterpreting the MNSQ statistics caused by the relatively small sample size of this study, the significance of MNSQ values were determined by the standardized fit statistic (ZSTD). $ZSTD > 2.0$ indicates that the corresponding MNSQ value is significant at $\alpha = 0.05$ (Linacre, 2019). Therefore, the misfit criteria in this study were determined by significant infit or outfit MNSQ values > 1.5 and ZSTD values for that infit or outfit MNSQ > 2.0 .

Previous studies have considered a scale built using Rasch analysis to be unidimensional when $< 5\%$ of the tasks used in the test or experiment fail to fit into the scale (Wright and Masters, 1982; Wright and Mok, 2000). For scales with < 20 tasks (inter-string distances), however, as is the case with our collision-avoidance experiment, a single misfit task would exceed this 5% criterion (see Hwang and Davies, 2009). Therefore, in this study, we set the criterion for unidimensionality as no more than one misfit inter-string distance within the scale. Finally, the task difficulties were expressed as logit. The logit of task difficulty was then used to determine whether the hierarchy of task difficulty was consistent with the expected difficulty of the inter-string distance, i.e. the smaller the inter-string distance, the more difficult the task. Rasch analysis was performed using MINISTEPS version 3.81.0 (www.winsteps.com/ministep.htm).

Reliability and validity of the collision-avoidance experiment

Reliability in Rasch analysis was reported as Cronbach's α , a coefficient of internal consistency that explains interrelatedness or inter-task correlations within the generated scale and indicates how well the tasks fit together as a scale (DeVon et al., 2007). The value of Cronbach's α ranges from 0 to 1 in which the higher values indicate a more reliable scale (DeVon et al., 2007). The species separation ratio and the task separation ratio, which indicate how well the dataset is spread out along the clutter negotiating ability scale of the experiment, were also reported. In our analysis, a species separation ratio was used to order and classify the bats' clutter negotiating ability and the tasks separation ratio verified the hierarchy of the inter-string distances. Species separation ratios less than 2, which give Cronbach's α values of less than 0.8 [where $\alpha = \text{separation ratio}^2 / (1 + \text{separation ratio}^2)$] are considered low, suggesting that the performance scale generated may not be sensitive enough to distinguish between high and low performers (Linacre, 2019). However, for task separation, ratio values less than 3 (Cronbach's α less than 0.9) are considered low, suggesting that the species sample is not large enough to confirm the task difficulty hierarchy. High separation ratios indicate that the clutter negotiating ability scale generated is valid, i.e. that it measures what it claims to measure (Messick, 1995; Borsboom et al., 2004).

Relationship between wing morphology and clutter negotiating ability

We calculated species means for body mass (M), wing dimensions (B , wing span; S , wing area; TL , wingtip length ratio; and TS , wingtip area ratio) and wing parameters (WL , wing loading; AR , wing aspect ratio; and I , wingtip shape index) for all 15 species. The body mass and wing dimensions (M , B , S , TL and TS) are primary measures in the study of animal flight morphology (Norberg and Rayner, 1987). The value of the wing aspect ratio and wing loading that describe, respectively, the size and shape of the wings were derived from body mass and the wing dimensions B and S (Eqns 4

and 5). The ratios of the wing-tip length (Eqn 6) and wing-tip area (Eqn 7) were derived from the wing dimensions TL and TS , and these ratios were used to determine the wingtip shape index (I) of an animal (Eqn 8). Analyses of body mass and wing dimension variables (M , B , S , TL and TS) were run separately from those of wing parameter variables (WL , AR and I). We tested the relationship between clutter negotiating ability and body mass and wing dimensions using simple linear regression. We then ran stepwise multiple regressions with clutter negotiating ability score (in logit unit scale) as the dependent variable and body mass and wing dimensions to determine which variables best explained variation in clutter negotiating ability among species. Next, we investigated the relationships between clutter negotiating ability and wing parameters following the same procedure: simple linear regressions and stepwise multiple regressions. The stepwise multiple regression analysis was based on stepwise variable selection or discriminant analysis algorithms in which the variable with the 'best' value for the criterion statistic is entered first. Analyses were performed in SPSS version 17.0 statistical packages for Windows (SPSS Inc., Chicago, IL, USA) and all data were log transformed before analysis.

RESULTS

Four-hundred and twenty-six individuals of 15 bat species that forage in the forest interior completed tasks in the experiment (Table S2). The individuals belong to three families: the Hipposideridae (*Hipposideros bicolor*, *Hipposideros cervinus*, *Hipposideros diadema*, *Hipposideros kunzi* and *Hipposideros ridleyi*); Rhinolophidae (*Rhinolophus affinis*, *Rhinolophus lepidus*, *Rhinolophus stheno* and *Rhinolophus trifolius*); and Vespertilionidae (*Kerivoula intermedia*, *Kerivoula papillosa*, *Kerivoula pellucida*, *Murina cyclotis*, *Murina suilla* and *Myotis ridleyi*). Morphological measurements were taken from the 426 individuals and an additional 335 individuals from the same 15 species that were not flown through the obstacle course (Table S3).

Clutter negotiating ability scale

We detected no misfits in the dataset, supporting unidimensionality of the clutter negotiating ability scale used in this study. Separation and reliability statistics for both species and tasks were high, with species separation of 3.60 and species reliability (Cronbach α) of 0.93, and inter-string distance separation for the dataset of 5.48 with a reliability value of Cronbach $\alpha = 0.97$. The high separation ratios and reliability statistics of our dataset demonstrated that the clutter negotiating scale generated was valid.

Species clutter negotiating ability and task difficulty appear along the same scale in the map determined by Rasch analysis, commonly referred to as Wright Map (or a Wright item-person map) (Fig. 3). Species appear in increasing order of clutter negotiating ability with species with poor performance in the collision-avoidance experiment at the bottom of the scale and species with greater clutter negotiating ability at the top of the scale (Fig. 3, left). Task difficulty (i.e. inter-string distance) appears in an inverse order compared with species clutter negotiating ability, with more difficult tasks (smaller inter-string distances) at the top and easier tasks (greater inter-string distances) at the bottom of the scale (Fig. 3, right). The sequence of the inter-string distance difficulty was in the expected order, with the smallest $D_{10\text{cm}}$ at the top of the scale and greatest $D_{60\text{cm}}$ at the bottom of the scale. Theoretically, the species clutter negotiating ability and task difficulty are comparable when species and tasks are aligned with each other on the map. Species that aligned with a specific task (inter-string

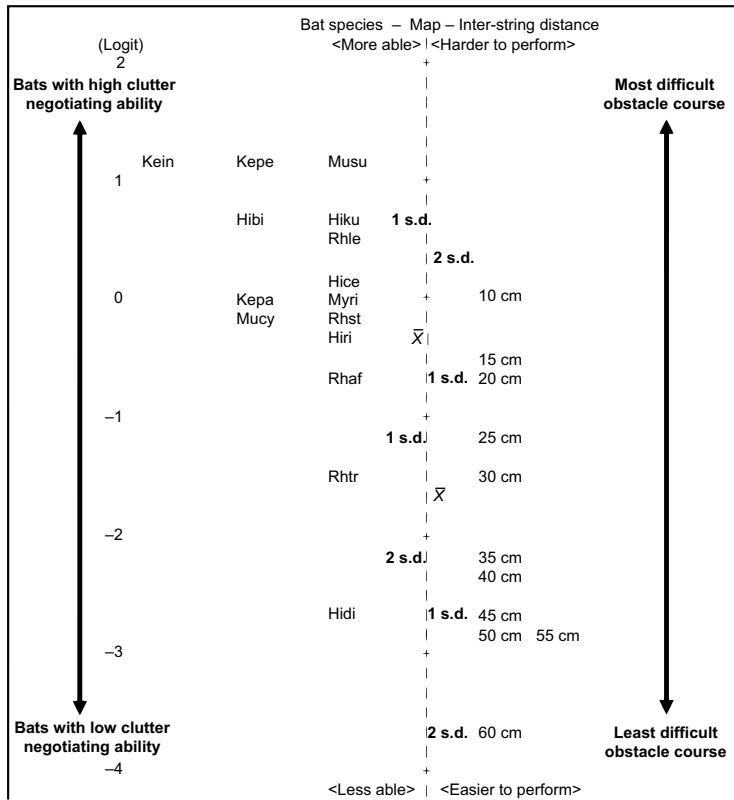


Fig. 3. Wright map presenting the hierarchical order of bat species clutter negotiating ability and inter-string distance difficulty. Logit values of the clutter negotiating ability given on the left. Clutter negotiating ability levels are represented as abbreviated species names and aligned to the left of the corresponding measure. Inter-string distances are aligned to the right of the corresponding values. \bar{X} , mean of species clutter negotiating ability or inter-string distance difficulty distribution; 1 or 2 s.d., one or two standard deviations from species clutter negotiating ability or inter-string distance distribution mean. Hibi, *Hipposideros bicolor*; Hice, *Hipposideros cervinus*; Hidi, *Hipposideros diadema*; Hiku, *Hipposideros kunzi*; Hiri, *Hipposideros ridleyi*; Kein, *Kerivoula intermedia*; Kepa, *Kerivoula papillosa*; Kepe, *Kerivoula pellucida*; Mucy, *Murina cyclotis*; Musu, *Murina suilla*; Myri, *Myotis ridleyi*; Rhaf, *Rhinolophus affinis*; Rhle, *Rhinolophus lepidus*; Rhst, *Rhinolophus stheno*; and Rhtr, *Rhinolophus trifoliatus*.

distance) have a 50% probability of clearing that inter-string distance (i.e. achieving a perfect score by flying through the task without breaking or touching any of the strings on all four banks of vertical strings). Five species aligned with specific inter-string distances in this study. For example, *R. trifoliatus* aligned with $D_{30\text{cm}}$, indicating a 50% probability that *R. trifoliatus* would clear that inter-string distance. By extension, *R. trifoliatus* has less than 50% chance of clearing inter-string distances located above the 30 cm task on the flight performance scale, but >50% chance of clearing inter-string distances below it. Similarly, *H. diadema* aligned with $D_{45\text{cm}}$, *R. affinis* aligned with $D_{20\text{cm}}$ and both *K. papillosa* and *M. ridleyi* were aligned with $D_{10\text{cm}}$. When species did not align with a specific task, the inter-string distance that would have a threshold of 50% performance is between the nearest measured inter-string distances.

The mean and two standard deviation points from the mean are shown for both species clutter negotiating ability (left-hand side of the map) and task difficulty (right-hand side) (Fig. 3). From the map, the mean (\bar{X}) of species clutter negotiating ability is approximately 1.5 s.d. above the mean (\bar{X}) of task difficulty. This suggests that bats with average clutter negotiating ability were able to perform tasks about 1.5 s.d. above the mean task difficulty, indicating that overall the test was easy to perform. The logit measures corresponding to both species clutter negotiating ability and inter-string distance difficulty depicted in Fig. 3 are reported in Table 1 and Table 2. The measurement range and coverage of the flight performance scale are illustrated on the left in Fig. 3, and cover 4.96 logits (between 1.16 and -3.80 logits).

When we compared the distribution of species clutter negotiating ability with that of task difficulty, seven species (*K. intermedia*,

Table 1. Inter-string distance difficulty statistics in measure order

Total score	Total count	Measure	Model	Infit		Outfit		Inter-string distance (cm)
				MNSQ	ZSTD	MNSQ	ZSTD	
122	15	-0.0518	0.0532	0.69	-0.6	0.56	-0.5	10
167	15	-0.5476	0.0817	1.38	0.7	1.56	1.0	15
175	15	-0.7164	0.0768	0.67	-0.4	1.03	0.3	20
194	15	-1.2171	0.1303	1.05	0.3	0.59	0.0	25
198	15	-1.4318	0.1164	0.24	-0.5	0.23	-0.4	30
211	15	-2.1811	0.1992	0.41	-0.6	0.10	-0.7	35
212	15	-2.3144	0.2134	0.01	-1.9	0.01	-1.2	40
217	15	-2.7232	0.1034	0.01	-2.9	0.02	-1.1	45
220	15	-2.8462	0.1396	0.00	-1.3	0.01	-1.2	50
220	15	-2.8462	0.1396	0.00	-1.3	0.01	-1.2	55
225	15	-3.7975	0.5049	Minimum measure				60

Inter-string distance is the distance between obstacles, i.e. the degree of clutter. Total score is the performance score for each inter-string distance across all species. Total count is the number of species that were tested for each inter-string distance. Measure and model were derived from Rash analysis. For details, see Materials and Methods. MNSQ, mean-square; ZSTD, standardized fit statistic.

Table 2. Bat clutter negotiating ability statistics in measure order

Total score	Total count	Measure	Model s.e	Infit		Outfit		Bat species	Cluster
				MNSQ	ZSTD	MNSQ	ZSTD		
165	11	1.1578	0.5211	Maximum measure				Kein	1
165	11	1.1578	0.5211	Maximum measure				Kepe	1
165	11	1.1578	0.5211	Maximum measure				Musu	1
163	11	0.6629	0.2239	0.11	-1.3	0.04	-0.9	Hiku	2
163	11	0.6629	0.2239	0.11	-1.3	0.04	-0.9	Hibi	2
162	11	0.5191	0.2037	0.20	-0.7	0.09	-0.7	Rhle	3
156	11	0.1033	0.0992	1.19	-0.5	0.46	0.0	Hice	4
154	11	0.0478	0.0899	0.33	-1.2	0.22	-0.4	Myri	4
152	11	-0.0032	0.0919	1.35	-0.7	0.47	0.0	Kepa	4
148	11	-0.1462	0.1249	0.33	-0.4	0.15	-0.5	Mucy	4
147	11	-0.1973	0.1288	0.50	-0.2	0.20	-0.4	Rhst	4
145	11	-0.2952	0.1168	2.33	1.2	1.03	0.5	Hiri	4
127	11	-0.7093	0.0917	0.14	-1.3	0.14	-0.5	Rhaf	4
99	11	-1.4707	0.0976	0.47	-0.5	0.61	-0.3	Rhtr	5
50	11	-2.6470	0.0676	0.14	-1.9	1.43	0.7	Hidi	5

Total score is the performance score for each species across all inter-string distances. Total count is the number of inter-string distances completed by each species.

Hibi, *Hipposideros bicolor*; Hice, *Hipposideros cervinus*; Hidi, *Hipposideros diadema*; Hiku, *Hipposideros kunzi*; Hiri, *Hipposideros ridleyi*; Kein, *Kerivoula intermedia*; Kepa, *Kerivoula papillosa*; Kepe, *Kerivoula pellucida*; Mucy, *Murina cyclotis*; Musu, *Murina suilla*; Myri, *Myotis ridleyi*; Rhaf, *Rhinolophus affinis*; Rhle, *Rhinolophus Lepidus*; Rhst, *Rhinolophus stheno*; and Rhtr, *Rhinolophus trifoliatius*.

Species in cluster 1 attained the maximum scores possible. WinSteps uses Bayesian logic to provide measures corresponding to those scores.

K. pellucida, *M. suilla*, *H. kunzi*, *H. bicolor*, *R. lepidus* and *H. cervinus*) were scaled above the most difficult inter-string distance, $D_{10\text{cm}}$, suggesting that these species were able to clear the most difficult task in the experiment with a probability above 50%. Indeed, three of these species (*K. intermedia*, *K. pellucida* and *M. suilla*) cleared all tasks and shared the maximum clutter negotiating ability at 1.16 logits (Table 2). Thus, the most difficult inter-string distance set for the experiment was not able to differentiate among the abilities of these three species, suggesting that smaller inter-string distances were needed to fully characterize their clutter negotiating ability and maybe improve discrimination. At the easy end of the inter-string distance scale, there were three inter-string distances ($D_{50\text{cm}}$, $D_{55\text{cm}}$, $D_{60\text{cm}}$) that had above 50% chance of being cleared by the poorest performer (*H. diadema*). All species were able to clear $D_{60\text{cm}}$, the minimum inter-string distance (Table 1). This suggests that the four widest inter-string distances were very easy tasks to perform, with $D_{60\text{cm}}$ not contributing to the discrimination among species at all. In general, the comparison between species clutter negotiating ability and task difficulty suggests that this collision-avoidance experiment was easy to perform for most of the bats tested.

There was little variability in scores of passes within a trial at inter-string distances above and below the 50% threshold. Around threshold, there was greater variability in pass scores within a trial and among individuals (see Fig. S1). Some individuals showed the anticipated pattern, with highest scores in the middle passes (e.g. some of the individuals of *K. papillosa* at $D_{15\text{cm}}$, *R. trifoliatius* at $D_{25\text{cm}}$ and *R. lepidus* at $D_{10\text{cm}}$ and $D_{15\text{cm}}$), but this was not consistent. Occasionally, individuals appeared to 'give up' despite initial success (see *R. trifoliatius* at $D_{15\text{cm}}$ and $D_{30\text{cm}}$) or had lowest scores in the middle passes (e.g. *R. trifoliatius* at $D_{25\text{cm}}$, *H. kunzi* at $D_{10\text{cm}}$).

Species clustering on the clutter negotiating ability scale

Examination of estimated true scores of species clutter negotiating ability and inter-string distance difficulty revealed breaks in connectivity within the dataset. Clusters (subsets) are not arbitrary but are identified by WinSteps based on breaks in connectivity when the data are modeled in a Graph Theory framework (Linacre, 2019). Rasch analysis of data subsets and connection reported that

the species fitted into the clutter negotiating ability scale in five different subsets or clusters (Table 2). The first cluster comprised species that had the highest clutter negotiating ability scores in negotiating all the inter-string distances (*K. intermedia*, *K. pellucida* and *M. suilla*). They were placed at the top of the clutter negotiating ability scale with 1.16 logits. The second cluster contained *H. kunzi* and *H. bicolor* at 0.66 logits. The third cluster was represented by a single species, *R. lepidus*, at 0.52 logits. Seven species made up the fourth cluster: *H. cervinus*, *M. ridleyi*, *K. papillosa*, *M. cyclotis*, *R. stheno*, *H. ridleyi* and *R. affinis* at 0.10, 0.05, 0.00, -0.15, -0.20, -0.30 and -0.71 logits, respectively. The fifth cluster was represented by two species of bats: *R. trifoliatius* and *H. diadema* at -1.47 and -2.65 logits, respectively.

Association between clutter negotiating ability and wing morphology

There were significant negative correlations between clutter negotiating ability and body mass (M), wing area (S) and wing span (B), and a positive correlation between clutter negotiating ability and wingtip length ratio (TL) and wingtip area ratio (TS) (Table 3). This indicates that species with smaller body mass, smaller wing area and smaller wing span but large wingtip area performed better in our collision-avoidance experiment. However, in the stepwise multiple regression analysis, a significant model emerged ($F_{1,14}=63.075$, $r^2=0.829$, $P<0.0001$) with body mass as the only significant predictor of clutter negotiating ability (Fig. S2). This suggests that clutter negotiating ability in the bats tested was primarily predicted by lighter body mass rather than the wing dimensions measured (Fig. 4). This likely also explains the significant positive relationship between wing loading (WL) and clutter negotiating ability (Fig. 5), as wing loading is in large part determined by (Eqn 5) and correlates tightly with body mass (Table 3).

DISCUSSION

We posed three questions. (1) Are collision-avoidance experiments reliable and a valid test of clutter negotiating ability? (2) If so, what is the relationship between clutter negotiating ability and wing morphology within the forest interior foraging ensemble? (3) Do

Table 3. Pearson correlation coefficient matrix of clutter negotiating ability, body mass and wing dimension (wing span, wing area, wingtip length ratio and wingtip area ratio) and wing parameters (wing loading, aspect ratio and wingtip shape)

	Clutter negotiating ability	<i>M</i>	<i>B</i>	<i>S</i>	<i>TL</i>	<i>TS</i>	<i>WL</i>	<i>AR</i>
<i>M</i>	-0.911**							
<i>B</i>	-0.863**	0.976**						
<i>S</i>	-0.854**	0.972**	0.992**					
<i>TL</i>	0.454	-0.608*	-0.544*	-0.490				
<i>TS</i>	0.591*	-0.677**	-0.586*	-0.531*	0.940**			
<i>WL</i>	-0.889**	0.918*	0.822*	0.799**	-0.727**	-0.835**		
<i>AR</i>	-0.317	0.313	0.351	0.232	-0.559*	-0.578*	0.409	
<i>I</i>	0.270	-0.051	0.015	-0.014	-0.476	-0.171	-0.111	0.229

M, body mass; *B*, wing span; *S*, wing area; *TL*, wingtip length ratio; *TS*, wingtip area ratio; *WL*, wing loading; *AR*, aspect ratio; *I*, wingtip shape index.

* $P < 0.05$ (two-tailed). ** $P < 0.01$ (two-tailed).

differences in wing morphology and clutter negotiating ability provide a potential mechanism for resource partitioning within the forest interior ensemble? Overall, we found that (1) collision-avoidance experiments were valid and reliable, and clutter negotiating ability differed among species from the forest interior ensemble and (2) scaled negatively with measures of size (i.e. body mass, wing area and wingspan) as well as wing loading. Body mass was a key determinant of performance. (3) Species within the ensemble could be grouped by performance into five different sub-ensembles, promoting differentiation of the forest interior foraging space.

Collision-avoidance experiments as reliable and valid tests of clutter negotiating ability

The scaling method used to analyze our dataset provided a true score estimation of clutter negotiating ability of bats in a collision-avoidance experiment that enabled us to evaluate species ability directly. Because of our approach, we could determine the reliability and validity of the experimental design and identify tasks that could

be modified in future experiments. In this study, the linear scale constructed by Rasch analysis allowed an easy comparison of inter-string distance difficulty and species clutter negotiating ability along the same scale, and indicated thresholds of ability, based on 50% probability of clearing a specific task, that allow discrimination among species. This is a powerful approach as it allows the relationship between species clutter negotiating ability and task difficulty to be evaluated simultaneously.

The clutter negotiation ability scale discrimination for species with maximum measures of ability beyond the most difficult inter-string distance (10 cm) was less informative. Although we know that the species (*K. intermedia*, *K. pellucida* and *Murina suilla*) could clear the most difficult inter-string distance, we do not know the limits of their ability. This illustrates poor targeting of inter-string distance difficulty to species ability and suggests that adding more challenging inter-string distances, e.g. 8, 6, 4 and 2 cm, might improve the targeting of inter-string distance difficulty to species clutter negotiating ability and discriminate among the top three species. For these three species, measures of clutter negotiating

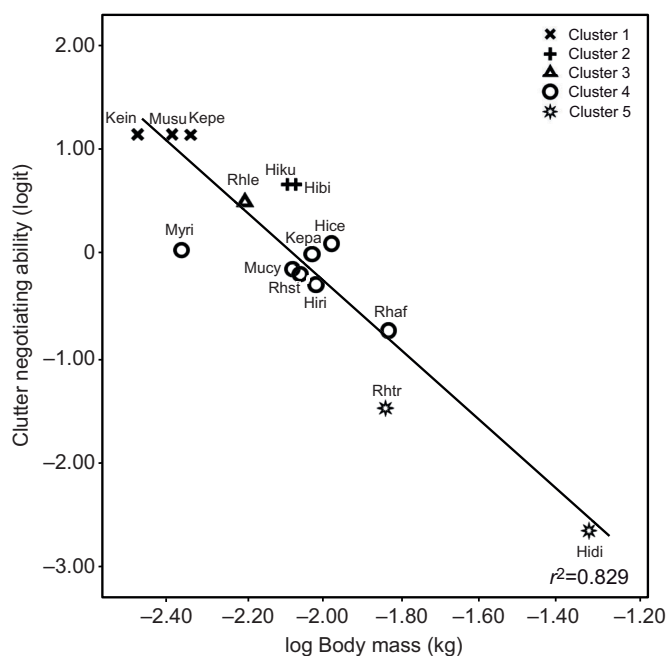


Fig. 4. Relationship between clutter negotiating ability and body mass across 15 species of insectivorous bats. Performance clusters are indicated by different symbols. Species abbreviations as in Fig. 3. $Y = -3.325X - 6.919$, $P < 0.0001$.

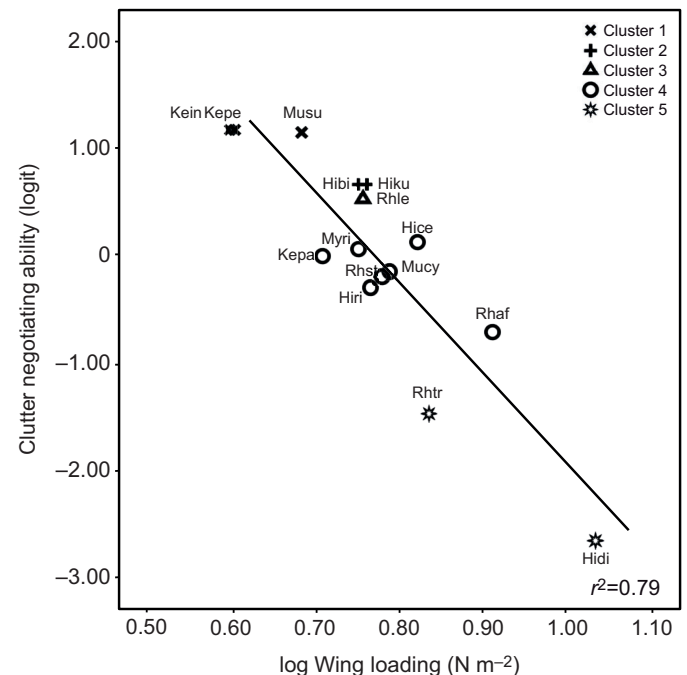


Fig. 5. Relationship between flight ability and wing loading across 15 species of insectivorous bats. Performance clusters are indicated by different symbols. Species abbreviations as in Fig. 3. $Y = -8.357X - 6.425$, $P < 0.0001$.

ability used to test relationships with morphology were generated using Bayesian logic (Lineacre, 2019). More challenging tasks to discriminate between the top three species could have provided measures that changed the linear relationships reported, so these relationships should be considered with care. In previous studies, the most challenging inter-string distance was set at half the wingspan of the smallest bat species tested. Stockwell (2001) tested five species of bats with obstacles spaced at multiples of 0.5, 0.75 or 1 of the bats' wingspan. For example, the obstacle distances for *Carollia castanea* with a 30 cm wingspan (the smallest wingspan among species tested) were set at 15, 22.5 and 30 cm. However, Aldridge (1986) set the smallest inter-string distance at an absolute value of 11 cm for *Myotis lucifugus* and *Myotis yumanensis* with wingspans of 25 and 24 cm, respectively. In both cases, the bats were highly successful in negotiating the smallest inter-string distance for the obstacle course (Aldridge, 1986; Stockwell, 2001), suggesting that they too failed to target inter-string distance difficulty to species clutter negotiating ability.

Relationship between clutter negotiating ability and wing morphology within the forest interior foraging ensemble

Although clutter negotiating ability significantly correlated with body mass, wingspan, wing area, wingtip area ratio and wing loading in our study, body mass and wing loading were the only significant predictors in the stepwise regressions. The importance of wing loading has been reported in other collision-avoidance experiments, although typically in combination with other parameters. A comparison of two *Myotis* species (Aldridge, 1986) found that both wing loading and wingspan played a significant role, and wing loading and wingtip shape explained performance differences between *Phoniscus papuensis* and *Nyctophilus bifax* (Rhodes, 1995). Jones et al. (1993) attributed differences in performance between *Hipposideros ruber* and *Asellia tridens* to differences in wing span, wing loading and aspect ratio. Aspect ratio was not a significant predictor of clutter negotiating ability in this study. Expectations of the role of aspect ratio in bat flight behavior come from Norberg and Rayner's (1987) original models, which treated bat wings as rigid and fixed. More recently, Schunk et al. (2017) found no influence of aspect ratio on coefficients of lift or drag in models of flapping compliant membrane wings and questioned the significance of aspect ratio as a predictor of bat flight behavior and ecology over the range of values seen in bats other than the narrow-winged molossid. Aspect ratio values in the present study were both low and of limited range (half-span values of 2.435–3.049 calculated from Table S3), falling within the lowest third of the aspect ratio distribution described in fig. 1 of Schunk et al. (2017). Interestingly, Stockwell (2001) reported that a species' overall size seemed to be a more important maneuverability criterion than specific wing morphology parameters for five phyllostomid bats.

In this study, the wingspan of the three species that had the maximum measures of clutter negotiating ability (*K. intermedia*, *K. pellucida* and *M. suilla*) were 21.5, 23.3 and 21.0 cm, respectively, close to the 0.5 ratio of other studies. The bat with the shortest wingspan in this study is *M. ridleyi* with 20.5 cm, a species that was not one of the highest performers, although it was still estimated to have 50% chance of clearing the 10 cm inter-string distance. Interestingly, *K. papillosa*, with a 30 cm wingspan, was also estimated to have 50% chance of clearing $D_{10\text{cm}}$. This suggests that wingspan is not reliable for scaling the inter-string distance in obstacle avoidance experiments, as it is not a reliable determinant of ability, and that future studies should select absolute inter-string

distances based on likely challenges faced by foraging bats of the study. As our morphological analyses illustrated, wing loading, not wingspan, was the best predictor of performance in this study. Whereas wing loading for *M. ridleyi* and *K. papillosa* was 5.609 N m^{-2} and 5.082 N m^{-2} , respectively, the three most able species, *K. intermedia*, *K. pellucida* and *M. suilla*, were characterized by extremely low wing loadings of 3.928, 3.963 and 4.830 N m^{-2} , respectively.

In our study, performance in the collision-avoidance experiment represented the overall latent trait 'clutter negotiating ability', but this encompasses diverse flight behaviors. Characterization of these behaviors was not our objective and would require multiple high-speed video cameras to capture the full range of wing and body kinematics used across 15 species. It would be particularly interesting to explore how species from different families with comparable clutter negotiating ability avoid obstacles. The experiment was conducted in natural darkness, but some observations of repeated behaviors included a tendency to hover at the start of the course in *Murina* species, and a slight gain in height before or while entering the course in most of the species tested, especially in some of the *Rhinolophus*. *Kerivoula pellucida* and *M. suilla* held both wings directly overhead when passing through small inter-string distances. In general, there were two paths selected by species, a mostly direct path that required the bat to 'slalom' through the offset strings and a diagonal path that exploited the string offset to allow straight flight. *Kerivoula pellucida* frequently completed the course less than 30 cm from the ground.

Our collision-avoidance course was designed to test the ability of forest interior bats to negotiate cluttered habitats. While flight ability was the focus, the course also presented a sensing and sensorimotor challenge, tasking bats to detect and assess the nature of the physical challenge before and during attempts to pass through. Paleotropical insectivorous bats that forage in cluttered habitats have sophisticated echolocation signal designs that enable them to detect prey in vegetatively dense habitats and avoid collisions while searching for and pursuing prey. The Kerivoulineae and Murinineae use short-duration, low-intensity, high-frequency calls that sweep rapidly through a broad band of frequencies (Kingston et al., 1999). This design allows them to detect prey very close to background (Schmieder et al., 2012) and likely separate clutter from prey echoes by segregating off-axis low-frequency echoes coming from the background from on-axis full-spectrum echoes that are focused on prey (Wohlgemuth et al., 2016). The Rhinolophidae and Hipposideridae use long narrowband calls of constant frequency (CF) matched to an acoustic fovea to detect the fluttering of insect prey wings (Schnitzler and Denzinger, 2011). Obstacles (clutter) are detected by analysis of differences in interaural sound intensity and travel time of returning echoes that are greatly accentuated by rapid alternation of ear positions (Mogdans et al., 1988; Vanderelst et al., 2015). Within these designs, species differ in echolocation frequency parameters, most notably the frequency of maximum energy (FMAXE) in the CF bats and bandwidth in the *Kerivoula* and *Murina*. Call parameter differences could potentially influence sensory perception and detection of the obstacle course resulting in performance differences that are independent of wing morphology. Moreover, if differences in sensory perception relate to body mass, this might confound our interpretation that body mass is influencing performance through the obstacle through flight.

We believe sensory constraints had minimal effect on performance because the string diameter, at 1.5 mm, should be readily detectable to all bat species in the study several meters from the first bank. Detection of small objects is determined in large part

by signal wavelength – the longer the wavelength (and hence lower the frequency), the more difficult it is to detect small objects. The bats with the lowest performance, *H. diadema* and *R. trivoliatus*, are CF bats using an FMAXE of 60 kHz and 53 kHz, respectively (Kingston et al., 2000). So the question arises, could they detect the course adequately using those frequencies? Berry et al. (2004) ensounded harp trap strings of 0.79 mm with a range of frequencies, and then calculated minimum detection distances. They found that 50 kHz calls gave minimum estimates of maximum detection distance of 2.5 m, with greater detection distances for higher calls. Our strings were twice the diameter, and our lowest-frequency bats produce very high intensity calls [presumed to be closer to 110–120 dB within 0.1 m of the bat's mouth, rather than the 95 dB used in Berry et al.'s (2004) models] so we believe they were able to detect the course in time to maneuver. This conclusion is supported by the high scores of *R. trivoliatus* down to 35 cm ($D_{35\text{cm}}$), and that of *H. diadema* at $D_{60\text{cm}}$ (Table S2). Note that although body mass scales inversely with FMAXE in CF bats (Jones, 1999), the relationship differs between families. Consequently, although the body mass of *H. diadema* is three times greater than that of *R. trivoliatus* (47 g versus 14 g, respectively), the *H. diadema* CF call is 10 kHz higher. If call frequency were an important determinant of performance in this study, *H. diadema* should have outperformed *R. trivoliatus*; instead, it was in the lowest clutter negotiating ability cluster.

By inspection of values in Schmieder et al. (2012), there was no discernible relationship between any call parameters, including bandwidth, and performance in the *Kerivoula* and *Murina*, nor with call parameters and body mass. Species in the top-performing cluster 1 used bandwidths of ~94 kHz (*M. suilla* and *K. intermedia*) and 155 kHz (*K. pellucida*), whereas *K. papillosa* and *M. cyclotis* in cluster 3 typically used 115 kHz bandwidth call (Schmieder et al., 2012).

In sum, the sensory task we set our bats appears to be well below threshold of the capabilities of our focal species, particularly as the course is static and does not require prey detection. However, species almost certainly differ in the mechanisms by which sensory information integrates with the motor functions governing the acoustic sensing system itself (e.g. head movements, pulse repetition rates, call duration), as well as those controlling body and wing position while flying through the obstacle course (sensorimotor transformation and integration; see Ulanovsky and Moss, 2008, for review). Differences in sensorimotor transformation might have influenced performance, but this could not be determined in our study.

Differences in wing morphology and clutter negotiating ability provide a potential mechanism for resource partitioning within the forest interior ensemble

The discrimination information from our clutter negotiating scale grouped the insectivorous bats into five clutter negotiating ability clusters. In most human studies using Rasch analysis, researchers try to avoid clusters in their dataset because it suggests that subjects are not drawn from a population with a normal distribution of ability across the distribution of task difficulties (Linacre, 2019). Clusters indicate that subjects are coming from populations with different abilities so cannot be compared based on their performance on the same test. In this study, however, clusters may have ecological meaning, suggesting the presence of groups (sub-ensembles) within the forest interior ensemble characterized by differential use of space within the habitat. Information regarding foraging strategies of insectivorous bats from the paleotropical region is rare, limiting the

interpretation of our findings. However, based on our observations, the two members of the least able cluster, *H. diadema* and *R. trivoliatus*, are perch hunters. Typically, these species hang from a perch and acoustically scan the airspace of small gaps in the forest understory and midstory for insects. Upon detection of prey, they sally forth to catch it and return to the perch. Although the overall structure of the forest understory or midstory in which the bats live may be cluttered, the degree of clutter in the actual foraging space is low. Other species are not known to use perches but hunt continuously on the wing (hawking). Hawking insects near or within vegetation requires both high maneuverability (Norberg, 1994) and clutter-tolerant echolocation (Siemers and Schnitzler, 2004). In tests of the sensory constraints on prey detection in an uncluttered environment, four of the more able species in our study (*K. intermedia*, *K. papillosa*, *M. cyclotis* and *M. suilla*) were able to detect and catch insects less than 6 cm away from a high echo-reflecting background with best performance as close as 2.5 cm away from the background (Schmieder et al., 2012). Moreover, detection performance was related to signal bandwidth – the greater the bandwidth, the closer the prey could be to a background and still be detected. In combination with Schmieder et al. (2012), our findings provide empirical support for the correlation or co-adaptation between wing morphology and echolocation signal design to particular foraging niches (Norberg, 1994).

Although 30 species of insectivorous bats in Krau Wildlife Reserve are considered forest interior species (Kingston et al., 2006), only 15 completed all tasks in our collision-avoidance experiment. Several species occurred too rarely to complete all tasks, particularly because resetting the inter-string distance took all day, so inter-string distance could not be adjusted to capitalize on rare captures. However, with the exception of *Nycteris tragata* (family Nycteridae) and *Megaderma spasma* (Megadermatidae), the missing species were members of the same subfamily or genus as those tested and most fell within the range of morphological parameters represented, so we believe that the inclusion of missing species would be unlikely to change the dynamics of the clutter negotiating scale by much. That said, by adding smaller inter-string distances and adding more species/genera, the clutter negotiating scale could be more fully tested. Moreover, explanation of the species distribution on the clutter negotiating scale could be improved with knowledge of foraging strategy and echolocation structure of the species.

Even with the 15 species, our experiment identified five clutter negotiating ability clusters that might reflect differential use of the forest interior habitat. Several niche dimensions that might facilitate resource partitioning in this species-rich ensemble from Krau Wildlife Reserve have been explored. Kingston et al. (1999) and Schmieder et al. (2012) focused on echolocation signal design in species of Kerivoulinae and Murininae. Interspecific differences in call parameters and bandwidth suggest a role for sensory niche partitioning among species. Senawi et al. (2015) demonstrated that differences in bite force among species provide a mechanism by which food resources may be partitioned. Furthermore, Kingston et al. (2000) tested the distribution of echolocation calls, body mass and wing dimensions of 15 species from the families Hipposideridae and Rhinolophidae. Although they found no evidence of deterministic separation of echolocation call frequency alone, overdispersion suggestive of niche differentiation was detected in multivariate space derived from both echolocation and wing parameters. In this study, we demonstrated that wing loading and body mass relate to flight negotiating ability in cluttered environments, facilitating ecological separation of species in this diverse ensemble.

Rasch analysis

Transformation of qualitative scales to allow quantitative analyses is widely used in assessments of human performance and ability. Rasch analysis belongs to a family of approaches based on Item Response Theory (IRT) that emphasize how individual responses to a particular test item are influenced by qualities of the individual and qualities of tests. IRT was originally developed by Lord (1953) and was first used to assess ability in the field of psychometrics. IRT is widely used to calibrate and evaluate items in tests or questionnaires to score subjects on their abilities, attitudes or other latent traits. For example, all major educational tests in the USA, such as the Graduate Record Examination (GRE) and Scholastic Aptitude Test (SAT), are developed and assessed using IRT-based techniques. This is because IRT-based techniques can significantly improve measurement accuracy and reliability while providing potentially significant reductions in assessment time and effort (Kingston and Dorans, 1984). In recent years, IRT-based models have become increasingly popular in health outcomes research (Hays et al., 2000; Duncan et al., 2003), quality-of-life research (Edelen and Reeve, 2007) and clinical research (Holman et al., 2003; Reise and Waller, 2009).

Assessments of animal ability or performance have relied heavily on quantitative measures because these lend themselves to conventional statistical approaches, e.g. determining sprint speed of reptiles using treadmills (Losos and Irschick, 1996) or hi-speed video recordings (Irschick, 2000), and maximum bite force capacity of vertebrates using force transducers (Herrel et al., 2005; Santana et al., 2010; Erickson et al., 2012; Senawi et al., 2015). Constraints in animal performance studies occur when there is no physical measurement instrument (such as a treadmill or transducer) available to measure performance or when dealing with more descriptive performances (qualitative data). Instead, scales derived from observer evaluations might be more applicable, with IRT providing opportunities for analytical frameworks that expand our statistical horizon in exploring and understanding animal performance and its relationship to morphology.

Acknowledgements

We thank the Economic Planning Unit, Department of Prime Minister, Malaysia for permission to conduct research (UPE: 40/200/19/2781) and Department of Wildlife and National Park Peninsular Malaysia for permission to work in Krau. We also acknowledge field assistance provided by Earthwatch volunteers 2010, students from Universiti Sains Malaysia and Southeast Asian Bat Conservation Research Unit 2011, and field station coordinator and trap movers: Zamiza Zainal, Aman Enya and Amri Aman.

Competing interests

The authors declare no competing or financial interests.

Author contributions

Conceptualization: J.S., T.K.; Methodology: J.S.; Formal analysis: J.S.; Investigation: J.S.; Data curation: J.S.; Writing - original draft: J.S.; Writing - review & editing: T.K.; Visualization: J.S.; Supervision: T.K.

Funding

This research was funded by Bat Conservation International, American Society of Mammalogist, IDEA WILD and Ministry of Higher Education, Malaysia to J.S. and by Texas Tech University and Earthwatch Institute to T.K.

Supplementary information

Supplementary information available online at <http://jeb.biologists.org/lookup/doi/10.1242/jeb.203950.supplemental>

References

- Aldridge, H. D. J. N. (1986). Manoeuvrability and ecological segregation in the little brown (*Myotis lucifugus*) and Yuma (*M. yumanensis*) bats (Chiroptera: Vespertilionidae). *Can. J. Zool.* **64**, 1878-1882. doi:10.1139/z86-280

- Aldridge, H. D. J. N. (1987). Turning flight of bats. *J. Exp. Biol.* **128**, 419-425.
- Aldridge, H. D. J. N. and Brigham, R. M. (1988). Load carrying and maneuverability in an insectivorous bat: a test of 5% "rule" of radio-telemetry. *J. Mammal.* **69**, 380-382. doi:10.2307/1381393
- Aldridge, H. D. J. N. and Rautenbach, I. L. (1987). Morphology, echolocation and resource partitioning in insectivorous bats. *J. Anim. Ecol.* **56**, 763-778. doi:10.2307/4947
- Allen, M. J. and Yen, W. M. (2002). *Introduction to Measurement Theory*. Illinois: Waveland Press.
- Andersen, E. B. (1977). Sufficient statistics and latent trait models. *Psychometrika* **42**, 69-81. doi:10.1007/BF02293746
- Andrich, D. (1978). Application of a psychometric rating model to ordered categories which are scored with successive integers. *Appl. Psychol. Meas.* **2**, 581-594. doi:10.1177/014662167800200413
- Andrich, D. (1988). *Rasch Model for Measurement*. Newbury Park, California: SAGE Publications Inc.
- Andrich, D. (1995a). Models for measurement, precision and the non-dichotomization of graded responses. *Psychometrika* **60**, 7-26. doi:10.1007/BF02294426
- Andrich, D. (1995b). Further remarks on non-dichotomization of graded responses. *Psychometrika* **60**, 37-46. doi:10.1007/BF02294428
- Bergou, A. J., Swartz, S. M., Vajdani, H., Riskin, D. K., Reimnitz, L., Taubin, G. and Breuer, K. S. (2015). Falling with style: bats perform complex aerial rotations by adjusting wing inertia. *PLoS Biol.* **13**, e1002297. doi:10.1371/journal.pbio.1002297
- Berry, N., O'Connor, W., Holderied, M. W. and Jones, G. (2004). Detection and avoidance of harp traps by echolocating bats. *Acta Chiropt.* **6**, 335-346. doi:10.3161/001.006.0211
- Bond, T. G. and Fox, C. M. (2015). *Applying the Rasch Model: Fundamental Measurement in the Human Sciences*. New York: Taylor and Francis Group.
- Boone, W. J. (2016). Rasch analysis for instrument development: why, when, and how? *CBE Life Sci. Educ.* **15**, rm4. doi:10.1187/cbe.16-04-0148
- Boone, W. J., Starver, J. R. and Yale, M. S. (2014). *Rash Analysis in the Human Sciences*. Springer.
- Borsboom, D., Mellenbergh, G. J. and Van Heerden, J. (2004). The concept of validity. *Psychol. Rev.* **111**, 1061-1071. doi:10.1037/0033-295X.111.4.1061
- DeVon, H. A., Block, M. E., Moyle-Wright, P., Ernst, D. M., Hayden, S. J. and Lazzara, D. J. (2007). A psychometric toolbox for testing validity and reliability. *J. Nurs. Scholarsh.* **39**, 155-164. doi:10.1111/j.1547-5069.2007.00161.x
- Duncan, P. W., Bode, R. K., Lai, S. M. and Perera, S. (2003). Rasch analysis of a new stroke-specific outcome scale: the stroke impact scale. *Arch. Phys. Med. Rehabil.* **84**, 950-963. doi:10.1016/S0003-9993(03)00035-2
- DWNP-DANCED (2001). *Krau Wildlife Reserve Management Plan*. Kuala Lumpur: Department of Wildlife and National Parks.
- Edelen, M. O. and Reeve, B. B. (2007). Applying item response theory (IRT) modeling to questionnaire development, evaluation and refinement. *Qual. Life Res.* **16**, 5-18. doi:10.1007/s11136-007-9198-0
- Engelhard, G. (2012). *Invariant Measurement: Using Rasch Models in the Social Behavioral and Health Sciences*. New York: Psychology Press.
- Erickson, G. M., Gignac, P. M., Steppan, S. J., Lappin, K. A., Vliet, K. A., Brueggen, J. D., Inouye, B. D., Kledzik, D. and Webb, G. J. W. (2012). Insights into the ecology and evolutionary success of crocodylians revealed through bite-force and tooth-pressure experimentation. *PLoS ONE* **7**, e31781. doi:10.1371/journal.pone.0031781
- Fenton, M. B. (1990). The foraging ecology and behavior of animal-eating bats. *Can. J. Zool.* **68**, 411-422. doi:10.1139/z90-061
- Francis, C. M. (1989). A comparison of mist-nets and two designs of harp-traps for capturing bats. *J. Mammal.* **70**, 865-870. doi:10.2307/1381730
- Hambleton, R. K. and Jones, R. W. (1993). An NCME instructional module on: comparison of classical test theory and item response theory and their applications to test development. *Educ. Meas.* **12**, 38-47. doi:10.1111/j.1745-3992.1993.tb00543.x
- Hays, R. D., Morales, L. S. and Reise, S. P. (2000). Item response theory and health outcomes measurement in the twenty-first century. *Med. Care* **38**, Suppl. 9, 1128-1142. doi:10.1097/00005650-200009002-00007
- Herrel, A., Podos, J., Huber, S. K. and Hendry, A. P. (2005). Evolution of bite force in Darwin's finches: a key role for head width. *J. Evol. Biol.* **18**, 669-675. doi:10.1111/j.1420-9101.2004.00857.x
- Hodgkison, R., Balding, B. T., Zubaid, A. and Kunz, T. H. (2004). Habitat structure, wing morphology, and the vertical stratification of Malaysian fruit bats (Megachiroptera: Pteropodidae). *J. Trop. Ecol.* **20**, 667-673. doi:10.1017/S0266467404001737
- Holman, R., Glas, C. A. W. and De Haan, R. J. (2003). Power analysis in randomized clinical trials based on item response theory. *Control. Clin. Trials* **24**, 390-410. doi:10.1016/S0197-2456(03)00061-8
- Hwang, J.-L. and Davies, P. L. (2009). Rasch analysis of the school function assessment provides additional evidence for the internal validity of the activity performance Scales. *Am. J. Occup. Ther.* **63**, 369-373. doi:10.5014/ajot.63.3.369

- Irschick, D. J.** (2000). Comparative and behavioral analyses of preferred speed: *Anolis* lizards as a model system. *Physiol. Biochem. Zool.* **73**, 428–437. doi:10.1086/317733
- Jones, G.** (1999). Scaling of echolocation call parameters in bats. *J. Exp. Biol.* **202**, 3359–3367.
- Jones, G., Morton, M., Hughes, P. M. and Budden, R. M.** (1993). Echolocation, flight morphology and foraging strategies of some West African hipposiderid bats. *J. Zool.* **230**, 385–400. doi:10.1111/j.1469-7998.1993.tb02691.x
- Kingston, N. M. and Dorans, N. J.** (1984). Item location effects and their implications for IRT equating and adaptive testing. *Appl. Psychol. Meas.* **8**, 147–154. doi:10.1177/014662168400800202
- Kingston, T., Jones, G., Akbar, Z. and Kunz, T. H.** (1999). Echolocation signal design in Kerivoulinae and Murininae (Chiroptera: Vespertilionidae) from Malaysia. *J. Zool.* **249**, 359–374. doi:10.1111/j.1469-7998.1999.tb00771.x
- Kingston, T., Francis, C. M., Zubaid, A. and Kunz, T. H.** (2003). Species richness in an insectivorous bat assemblage from Malaysia. *J. Trop. Ecol.* **19**, 67–79. doi:10.1017/s0266467403003080
- Kingston, T., Jones, G., Zubaid, A. and Kunz, T. H.** (2000). Resource partitioning in rhinolophid bats revisited. *Oecologia* **124**, 332–342. doi:10.1007/PL00008866
- Kingston, T., Lim, B. L. and Zubaid, A.** (2006). *Bats of Krau Wildlife Reserve*. Bangi, Malaysia: Penerbit Universiti Kebangsaan Malaysia.
- Linacre, J. M.** (2019). *Winsteps® Rasch Measurement Computer Program User's Guide*. Beaverton, Oregon: Winsteps.com.
- Lord, F. M.** (1953). The relation of test score to the trait underlying the test. *Educ. Psychol. Meas.* **13**, 517–548. doi:10.1177/001316445301300401
- Losos, J. B. and Irschick, D. J.** (1996). The effect of perch diameter on escape behaviour of *Anolis* lizards: laboratory predictions and field tests. *Anim. Behav.* **51**, 593–602. doi:10.1006/anbe.1996.0063
- McKenzie, N. L., Gunnell, A. C., Yani, M. and Williams, M. R.** (1995). Correspondence between flight morphology and foraging ecology in some Palaeotropical bats. *Aust. J. Zool.* **43**, 241–257. doi:10.1071/ZO9950241
- Messick, S.** (1995). Validity of psychological assessment: validation of inferences from persons responses and performances as scientific inquiry into score meaning. *Am. Psychol.* **50**, 741–749. doi:10.1037/0003-066X.50.9.741
- Mogdans, J., Ostwald, J. and Schnitzler, H. U.** (1988). The role of pinna movement for the localization of vertical and horizontal wire obstacles in the Greater Horseshoe Bat, *Rhinolopus ferrumequinum*. *J. Acoust. Soc. Am.* **84**, 1676–1679. doi:10.1121/1.397183
- Norberg, U. M.** (1990). *Vertebrate Flight: Flight Mechanics, Physiology, Morphology, Ecology and Evolution*. Berlin: Springer-Verlag.
- Norberg, U. M.** (1994). Wing design, flight performance and habitat use in bats. In *Ecological Morphology: Integrative Organismal Biology* (ed. P. C. Wainright and M. Reilly), pp. 205–239. Chicago, Illinois: University of Chicago Press.
- Norberg, U. M.** (1998). Morphological adaptations for flight in bats. In *Bat Biology and Conservation* (ed. T. H. Kunz and P. A. Racey), pp. 93–108. Washington DC: Smithsonian Institution Press.
- Norberg, U. M. and Rayner, J. M. V.** (1987). Ecological morphology and flight in bats (Mammalia; Chiroptera): wing adaptations, flight performance, foraging strategy and echolocation. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **316**, 335–427. doi:10.1098/rstb.1987.0030
- Reise, S. P. and Waller, N. G.** (2009). Item response theory and clinical measurement. *Annu. Rev. Clin. Psychol.* **5**, 27–48. doi:10.1146/annurev.clinpsy.032408.153553
- Rhodes, M. P.** (1995). Wing morphology and flight behaviour of the golden-tipped bat, *Phoniscus papuensis* (Dobson) (Chiroptera: Vespertilionidae). *Aust. J. Zool.* **43**, 657–663. doi:10.1071/ZO9950657
- Salzberger, T.** (2010). Does the Rasch model convert an ordinal scale into an interval scale? *Rasch Measurement Transactions* **24**, 1273–1275. <https://www.rasch.org/rmt/rmt242a.htm>.
- Santana, S. E., Dumont, E. R. and Davis, J. L.** (2010). Mechanics of bite force production and its relationship to diet in bats. *Funct. Ecol.* **24**, 776–784. doi:10.1111/j.1365-2435.2010.01703.x
- Saunders, M. B. and Barclay, R. M. R.** (1992). Ecomorphology of insectivorous bats: a test of predictions using two morphologically similar species. *Ecology* **73**, 1335–1345. doi:10.2307/1940680
- Schnitzler, D. A., Kingston, T., Hashim, R. and Siemers, B. M.** (2012). Sensory constraints on prey detection performance in an ensemble of vespertilionid understorey rain forest bats. *Funct. Ecol.* **26**, 1043–1053. doi:10.1111/j.1365-2435.2012.02024.x
- Schnitzler, H. U. and Denzinger, A.** (2011). Auditory fovea and Doppler shift compensation: adaptations for flutter detection in echolocating bats using CF-FM signals. *J. Comp. Physiol. A Neuroethol. Sens. Neural Behav. Physiol.* **197**, 541–559. doi:10.1007/s00359-010-0569-6
- Schnitzler, H.-U. and Kalko, K. M. V.** (1998). How echolocating bats search and find food. In *Bat Biology and Conservation* (ed. T. H. Kunz and P. A. Racey), pp. 183–196. Smithsonian Institution Press.
- Schunk, C., Swartz, S. M. and Breuer, K. S.** (2017). The influence of aspect ratio and stroke pattern on force generation of a bat-inspired membrane wing. *Interface Focus* **7**, 20160083. doi:10.1098/rsfs.2016.0083
- Senawi, J., Schmieder, D., Siemers, B. and Kingston, T.** (2015). Beyond size - morphological predictors of bite force in a diverse insectivorous bat assemblage from Malaysia. *Funct. Ecol.* **29**, 1411–1420. doi:10.1111/1365-2435.12447
- Siemers, B. M. and Schnitzler, H. U.** (2004). Echolocation signals reflect niche differentiation in five sympatric congeneric bat species. *Nature* **429**, 657–661. doi:10.1038/nature02547
- Stockwell, E. F.** (2001). Morphology and flight manoeuvrability in New World leaf-nosed bats (Chiroptera: Phyllostomidae). *Zool. Soc. Lond.* **254**, 505–514. doi:10.1017/S0952836901001005
- Swartz, S. M., Iriarte-Díaz, J., Riskin, D. K. and Breuer, K. S.** (2012). A bird? A plane? No, it's a bat: an introduction to the biomechanics of bat flight. In *Evolutionary History of Bats: Fossils, Molecules and Morphology* (ed. G. F. Gunnell and N. B. Simmons), pp. 317–352. Cambridge University Press.
- Trochim, W.** (2006). *The Research Method Knowledge Base*. Cincinnati, Ohio: Atomic Dog Publishing.
- Vanderelst, D., Holderied, M. W. and Peremans, H.** (2015). Sensorimotor model of obstacle avoidance in echolocating bats. *PLoS Comput. Biol.* **11**, e1004484. doi:10.1371/journal.pcbi.1004484
- Ulanovsky, N. and Moss, C. F.** (2008). What the bat's voice tells the bat's brain. *Proc. Natl Acad. Sci. USA* **105**, 8491–8498. doi:10.1073/pnas.0703550105
- Wohlgemuth, M. J., Lou, J. and Moss, C. F.** (2016). Three-dimensional auditory localization in the echolocating bat. *Curr. Opin. Neurobiol.* **41**, 78–86
- Wright, B. D. and Masters, G. N.** (1982). *Rating Scale Analysis*. Chicago: MESA Press.
- Wright, B. D. and Mok, M.** (2000). Rasch model overview. *J. Appl. Meas.* **1**, 83–106.
- Wright, B. D. and Stone, M. H.** (1979). *Best Test Design*. Chicago: MESA Press.