Anim Cogn (2016) 19:329–342
DOI 10.1007/s10071-015-0935-4

ORIGINAL PAPER

Temporal averaging across multiple response options: insight into the mechanisms underlying integration

Benjamin J. De Corte[1] · **Matthew S. Matell**[1]

Received: 10 June 2015 / Revised: 16 October 2015 / Accepted: 22 October 2015 / Published online: 31 October 2015

Abstract Rats trained on a dual-duration, dual-modality peak-interval procedure (e.g., tone = 10 s/light = 20 s) often show unimodal response distributions with peaks that fall in between the anchor durations when both cues are presented as a simultaneous compound. Two hypotheses can explain this finding. According to the averaging hypothesis, rats integrate the anchor durations into an average during compound trials, with each duration being weighted by its respective reinforcement probability. According to the simultaneous temporal processing hypothesis, rats time both durations veridically and simultaneously during compound trials and respond continuously across both durations, thereby producing a unimodal response distribution with a peak falling in between the anchor durations. In the present compounding experiment, rats were trained to associate a tone and light with two different durations (e.g., 5 and 20 s, respectively). However, in contrast to previous experiments, each cue was also associated with a distinct response requirement (e.g., left nosepoke for tone/right nosepoke for light). On the majority of compound trials, responding on a given nosepoke fell close to its respective duration, but was shifted in the direction of the other cue's duration, suggesting rats timed an average of the two durations. However, more weight appeared to be given to the duration associated with the manipulandum on which the rat responded, rather than the duration associated with a higher reinforcement probability as predicted by the averaging hypothesis. Group differences were also observed, with rats trained to associate the tone and light with the short and long durations, respectively, being more likely to show these shifts than the counterbalanced modality–duration group (i.e., light-short/tone-long). This parallels group differences observed in past studies and suggest that cue weighting in response to stimulus compounds is influenced by the modality–duration relationship of the anchor cues. The current results suggest that temporal averaging is a more flexible process than previously theorized and provide novel insight into the mechanisms that affect cue weighting.

Keywords Interval timing · Time perception · Temporal averaging · Simultaneous temporal processing · Stimulus compounding

Introduction

Interval timing, timing in the seconds to minutes range, is crucial for a variety of processes such as optimal foraging (Kacelnik et al. 1990) and associative learning (Gibbon and Balsam 1981; Balsam and Gallistel 2009). Thus, to function effectively, it is necessary to have an accurate sense of the passage of time.

In animals, interval timing is frequently examined using the peak-interval (PI) procedure (Roberts 1981; Gallistel et al. 2004; Macdonald et al. 2012). In a typical PI task, the onset of a stimulus (e.g., a tone) indicates that reinforcement can be earned for making an operant response after a fixed "criterion duration" has elapsed (e.g., 10 s), without penalty for early responding. These are referred to as fixed-interval (FI) trials. Probe trials are also included, wherein the stimulus is presented for 3–4 times the length of the criterion duration and no reinforcement is provided. When

✉ Matthew S. Matell
Matthew.matell@villanova.edu

[1] Department of Psychology, Villanova University, 800 E. Lancaster Ave., Villanova, PA, USA

plotted as a function of time, the average response rate during probe trials typically resembles a Gaussian distribution, with a mode (i.e., "peak") centered over the time at which reward is usually earned. A considerable amount of work has shown that the spread of these distributions (i.e., the standard deviation) varies in proportion to the duration being timed (Gibbon et al. 1984; Gibbon and Church 1990). For example, if the criterion duration were doubled (e.g., from 10 to 20 s), then both the spread and the mode of the original response distribution would increase by a factor of two. This is referred to as scalar variance and is a hallmark of interval timing (Gibbon 1977).

Recently, we conducted a PI study in which rats were trained to associate two cues (tone and light) with two different durations (e.g., 10 and 20 s, respectively) (Swanton et al. 2009). In addition to trials in which each cue was presented alone, we also included "compound" probe trials in which both the tone and light were presented simultaneously. The primary purpose of introducing these trials was to assess how rats would respond when presented with cues that provided conflicting temporal information regarding when reward would become available. When cues were presented in isolation, responses were centered over the appropriate criterion times. In contrast, during compound trials, rats produced a unimodal, scalar response distribution with a peak time that fell in between the two criterion durations. Given that these distributions were scalar, we concluded that rats were timing a single temporal expectation that somehow reflected a combination of the two durations. Therefore, we referred to this effect as "temporal averaging."

Further work has gone on to investigate the nature of this averaging process (Swanton and Matell 2011; Kurti et al. 2013; Matell and Kurti 2014). For example, Swanton and Matell (2011) tested the effects of using different durations and duration ratios associated with the component cues (e.g., 8/24 s, 5/30 s, etc.) on responding during compound trials. Then, they evaluated how adequately different Pythagorean means (i.e., arithmetic, geometric, and harmonic) were able to account for the observed compound peak times. When temporal averaging occurred, peak times were predicted best by a geometric mean of the individual durations (i.e., an arithmetic mean with the durations being represented on a log scale). Furthermore, this fit was substantially improved when the peak time for each component cue was weighted by its respective reinforcement probability (reinforcement probability = [number of FI trials]/[number of probe trials + number of FI trials] for a given cue).

Therefore, there is support for the notion that rats are engaging in averaging behavior during compound trials, which we will refer to here as the "averaging hypothesis." However, colleagues and reviewers often offer an alternative explanation of compound responding when discussing this effect. Specifically, it has been suggested that rats may be timing both durations in a veridical manner during compound trials. If this is the case and the ratio between the two durations is small, then averaging across trials during analysis would create the appearance of a unimodal compound response distribution with a peak time that falls in between the two criterion durations. However, this unimodal peak would be an artifact of averaging across two different response patterns during analysis, and not the result of timing an intermediate duration. In the present article, we will refer to this explanation of compound responding as the "simultaneous temporal processing hypothesis."

This latter interpretation is consistent with prior work where animals have been explicitly trained to time multiple durations simultaneously (Meck and Church 1984; Olton et al. 1988; Pang et al. 2001). For example, Leak and Gibbon (1995) trained pigeons that the onset of a houselight signified that reward could be earned for responding on a single key after one of two durations had elapsed, determined randomly on each trial. When the ratio between the two durations was large (e.g., 1:4), response rates were bimodal with maximums falling over either criterion duration. However, at smaller ratios (e.g., 1:2.5), response distributions were broad and unimodal. In either case, pigeons were assumed to be timing both durations simultaneously. However, at smaller ratios, the decision to start responding at the longer duration frequently occurred before pigeons stopped responding at the short duration, causing them to respond continuously throughout the trial.

While the simultaneous temporal processing (STP) hypothesis can account for some patterns of compound responding, it cannot account for two aspects of prior compounding data. First, we have examined compound responding when the ratio between the component durations is large (e.g., 1:4, 1:6) (Swanton and Matell 2011; Kurti et al. 2013). According to the STP hypothesis, compound responding should have been bimodal in this case, similar to what Leak and Gibbon (1995) observed at larger ratios in their study. However, rats in these experiments showed unimodal compound response distributions falling in between the component peak times. Second, as this hypothesis predicts that rats should respond continuously across both component durations during compound trials, their response distributions should be broad and non-scalar. Contrary to this prediction, compound responding has been scalar in all prior studies (Swanton et al. 2009; Swanton and Matell 2011; Kurti et al. 2013; Matell and Kurti 2014).

In the current study, we sought to more explicitly evaluate the STP and averaging hypotheses. Similar to prior investigations, we trained rats to associate a tone and

light with two different durations and then introduced compound probe trials (e.g., Swanton et al. 2009). However, we also associated each cue with a distinct response requirement (i.e., the left or right nosepoke), such that reinforcement was delivered only if rats responded on a given cue's associated response manipulandum during a trial (e.g., tone-respond left/light-respond right). This contrasts to previous investigations, in which both cues have been associated with the same response requirement.

This design is similar, in principle, to a variant of the PI procedure referred to as the "bipeak" procedure (e.g., Meck et al. 2012). In this task, animals are trained that a single cue signifies reward availability after one of two different durations elapses, determined randomly on different trials. Unlike the design used by Leak and Gibbon (1995), these durations are also associated with distinct response requirements, such that responses on a given manipulandum are only reinforced if its respective duration has elapsed. During probe trials in this task, animals begin responding on the manipulandum associated with the shorter duration and subsequently switch to responding on the manipulandum associated with the longer duration. Consequently, distinct response distributions are formed at both manipulanda with peak times centered over their respective criterion durations.

The averaging and STP hypotheses make divergent predictions regarding how responding should proceed when the compound cue is presented during the current study. If rats time both durations veridically during compound trials as suggested by the STP hypothesis, compound responding should resemble that seen during probe trials in the bipeak procedure. Specifically, rats should progress within a trial from responding on the short to the long cue's associated manipulandum. Importantly, distinct response distributions should form at both apertures, with peak times centered over their respective criterion durations. In contrast, according to the averaging hypothesis, rats should form an integrated temporal expectation that reflects a probability-weighted average of the two anchor durations during compound trials. Rats should use this expectation to guide responding on either aperture, and responses should randomly alternate between nosepokes (either within or between trials). Consequently, response distributions at both apertures should overlap.

Method

Subjects and apparatus

Subjects were 20 male Sprague–Dawley rats (Rattus norvegicus; Harlan, Indianapolis, IN) who were approximately 3 months of age at the start of the experiment. Rats were given ad libitum access to water but were placed on a restricted diet in order to maintain their body weights at 85–90 % of free-feed levels, adjusted for growth. Rats were housed in pairs, and colony room lights were set to a 12-h light–dark cycle. All training sessions occurred during the light cycle and took place in standard operant-conditioning chambers (30.5 × 25.4 × 30.5 cm; Coulbourn Instruments, Allentown, PA). The chambers' left and right sides consisted of ventilated Plexiglas. The front, back, and top sides of the chambers were composed of aluminum, and chamber floors consisted of an array of stainless steel bars. Three nosepoke apertures, equipped with photobeam detection circuits, were located on the back wall. An 11-lux houselight and seven-tone audio generator, set to produce 95-dB tones, were located along the back wall of the chambers as well. A pellet dispenser, located on the front wall, was used to deliver 45-mg grain pellets into a food magazine (Bio-Serv, Flemington, NJ). A continuously running fan operating at 60 dB provided ventilation. Stimulus control and data acquisition were accomplished using a standard operant-conditioning control program (Graphic State, Coulbourn Instruments, Allentown, PA), which had a temporal resolution of 20 ms.

Procedure

The experiment consisted of four stages: nosepoke training, fixed-interval training, peak-interval training, and compound testing. In the latter three stages, a 4-kHz tone and 11-lux houselight were used as discriminative stimuli. Each stimulus was associated with either a 5- or 20-s duration. For simplicity, we will commonly refer to these as the "short" and "long" durations, respectively. For half of the rats, the tone and houselight signaled reinforcement availability after the short and long durations elapsed, respectively (hereafter referred to as the tone-short/light-long or $T_S L_L$ group). For the remaining half, this modality–duration relationship was reversed (hereafter referred to as the light-short/tone-long or $L_S T_L$ group). Two days before training began, rats were given twenty 45-mg grain pellets in their home cage in order to acclimate them to the food used for reinforcement. Thereafter, rats were trained 5 days per week at approximately the same time each day. Sessions lasted 2 h.

Nosepoke training (four sessions)

During nosepoke training, trials began with the illumination of one of the three nosepokes (determined randomly at the start of each trial). An insertion of the snout into the illuminated nosepoke aperture resulted in reinforcement delivery and the initiation of a dark, uniformly distributed 2- to 8-s inter-trial interval (ITI). Responses during the ITI

or to non-illuminated nosepokes were not reinforced. This process repeated throughout the 2-h session.

Fixed-interval training (12 sessions)

During fixed-interval (FI) training, trials began with the onset of either the tone or houselight (randomly selected at the start of each trial). Reinforcement could be earned for making a nosepoke response after the current cue's respective duration had elapsed, without any programmed consequence for early responses. These durations were 5 and 20 s, and were associated with the tone and light in a counter-balanced manner, as described above. Importantly, reinforcement for each cue was dependent on responding at a distinct nosepoke. For half of the rats, the short and long cues were associated with the left and right nosepokes, respectively. The other half received the counterbalanced duration–nosepoke pairing. Combined with the modality of the cues, this resulted in four groups (tone-short-left/light-long-right, tone-short-right/light-long-left, light-short-left/-tone-long-right, light-short-right/tone-long-left). Trials terminated after reward delivery and were followed by a uniformly distributed, 60- to 80-s ITI.

Peak-interval training (24 sessions)

Peak-interval (PI) training was identical to FI training with one exception. In addition to trials in which reinforcement was delivered, probe trials were also introduced, during which cues were presented for 3–4 times the long duration (i.e., 60–80 s) and no reinforcement was provided. Reinforcement probabilities were initially set to 80 % for both durations (12 sessions). To generate equivalent response rates, the reinforcement probability for the short duration was then decreased in 20 % increments across successive sessions, while the reinforcement probability for the long duration remained constant. This continued until response rates were approximately equal for both cues (five sessions at 60 % short and seven sessions at 40 % short).

Compound testing (five sessions)

Compound testing was identical to PI training, except that compound probe trials were also introduced (20 % of all trials within a session). During these trials, both the tone and houselight were presented simultaneously. Like single-cue probe trials, compound probe trials lasted 60–80 s (uniformly distributed) and no reinforcement was delivered.

Analysis

All data presented below came from probe trials during the compound testing phase. Each entry into the nosepoke aperture was classified as a response, and another response was only recorded if the rat first removed and then re-inserted its snout into the nosepoke.

Peak functions

Each rat's responses were grouped into 1-s bins and averaged across trials to create peak functions (i.e., mean response rate as a function of time within a trial) for each cue. These peak functions were fit (curve fitting package of MATLAB, Cambridge, MA) using a five-parameter Gaussian function: $Y = Y_0 + A \times \exp((-1/2) \times (\mathrm{abs}(X - B)/C)\text{^}D)$, as used previously (Swanton et al. 2009). Y_0 represents the baseline rate; A is a scaling factor; B is the mean; C is the standard deviation; and D allows the function to accurately fit peaks with different levels of kurtosis. B and C were used as measures of peak time and spread, respectively.

Single-trials analysis

Previous work has demonstrated that behavior during probe trials follows a break-run-break pattern of responding (i.e., low response rate, high response rate, low response rate) (Church et al. 1994; Gibbon and Church 1990; Swanton et al. 2009). The times at which rats initiate and terminate the high state of responding, referred to as "start" and "stop" times, respectively, are known to be temporally controlled. Therefore, for each trial, we iteratively fit three flat lines (first low state, high state, and second low state), with the transition at each time bin, until the absolute residuals between these lines and the data were minimized (for a recent alternative method for this analysis, see Harris 2015). Start and stop times could not be reliably identified during trials where rats made less than three responses (7 % of trials in total). Therefore, these trials were not analyzed. From this analysis, we were able to identify the times at which rats started and stopped a high state of responding. Occasionally, some rats showed burst responding quite late during probe trials. These responses were rare and did not appear to be temporally controlled as seen previously (Matell et al. 2006). To prevent these responses from biasing the statistics describing the start and stop times, we excluded trials in which the start times occurred after 40 s (i.e., twice the length of the long criterion duration). This resulted in the exclusion of 7 % of trials in the $T_S L_L$ group and 5 % of trials in the $L_S T_L$ group.

Compound response pattern analysis

As mentioned previously, the averaging and simultaneous temporal processing hypotheses make different predictions regarding which nosepokes rats will respond on during

compound trials. To evaluate these predictions, we categorized the response patterns seen at each nosepoke during these trials using the data obtained from the single-trials analysis. Trials where rats responded solely on the short or long nosepokes were categorized as "short-only" and "long-only", respectively. If responses occurred on both nosepokes, the trial was classified based on the temporal order of the start and stop times at each nosepoke. For example, "bipeak" was used as a label for trials where the start and stop times on the short duration's nosepoke occurred before those on the long duration's nosepoke. In other words, rats progressed in time from responding on the short to the long duration's nosepoke.

This analysis revealed that 98 % of trials fell under the categories of short-only, long-only, or bipeak. The remaining 2 %, which we collectively refer to as "other", were mainly composed of trials where rats switched back and forth between nosepokes (e.g., start and stop times on the short duration's nosepoke fell between those of the long duration's nosepoke). We excluded this category from our dataset for two reasons. First, as these trials made up only 2 % of the data, it allowed us to focus on the more prominent response types. Second, and more importantly, these proportions were equal to zero for several of the rats (i.e., many rats did not engage in these response types), which would be problematic for the type of analysis used to examine the data (described below) (Martín-Fernández et al. 2003). Therefore, the remaining three proportions for each rat were normalized by their sum, such that the data once again added to one.

As these were proportional data, using conventional statistics would have been inappropriate for a variety of reasons. For example, each datum was not free to vary, as an increase in one proportion would necessarily cause a complementary decrease in the other two proportions. Therefore, we employed a technique known as compositional data analysis, which is specifically tailored for evaluating proportional data (Aitchison 1982, 2005). Briefly, this entails examining how each proportion relates to the others as opposed to evaluating each individually. Furthermore, placing these relations on a log-ratio scale allows the data to be subject to multivariate analysis (Pawlowsky-Glahn and Egozcue 2006). Therefore, we evaluated the data using isometric log-ratio (ILR) balances (for description, see Egozcue and Pawlowsky-Glahn 2005; Engle and Rowan 2013).

The first balance (B_1) allowed us to assess between-group differences in bipeak responding and is given by:

$$B_1 = \sqrt{\frac{2}{3}}\log\left(\frac{\sqrt{P_{\text{short-only}} \times P_{\text{long-only}}}}{P_{\text{bi-peak}}}\right) \quad (1)$$

where $P_{\text{short-only}}$ and $P_{\text{long-only}}$ refer to proportions of short-only and long-only responses, respectively, and P_{bipeak} refers to the proportion of bipeak responding. The resulting value reflects the extent to which responding on a single nosepoke (short-only or long-only) is seen relative to responding on both nosepokes (i.e., bipeak responding).

The second ILR balance (B_2) allowed us to assess between-group differences in short-only relative to long-only responding and is given by:

$$B_2 = \frac{1}{\sqrt{2}}\log\left(\sqrt{\frac{P_{\text{short-only}}}{P_{\text{long-only}}}}\right) \quad (2)$$

Statistics

Peak times and normalized spreads (i.e., coefficient of variation = peak spread/peak time) obtained from the fitted peak functions, as well as start and stop times obtained from the single-trials analysis, were analyzed individually using mixed-model analyses of variance (ANOVA). Group (T_SL_L, L_ST_L) and Nosepoke Pairing (short-left/long-right versus short-right/long-left) were used as between-subjects factors. Duration (short versus long, indicating from which nosepoke the data were obtained) and Compounding (Cue-Alone, Compound-Single-Nosepoke, and Compound-Bipeak) were used as within-subjects factors. "Cue-Alone" refers to trials in which either the short or long cue was presented in isolation. "Compound-Single-Nosepoke" refers to compound trials in which a rat responded on only one of the two nosepokes (short or long). "Compound-Bipeak" refers to compound trials in which a rat responded first on the short nosepoke and then switched to responding on the long nosepoke. When significant interactions were found, data from each group were analyzed separately, using paired t tests. ILR balances were examined using a multivariate analysis of variance (MANOVA), with Group and Nosepoke Pairing serving as fixed factors. A significance level of $\alpha = 0.05$ was used for all statistics reported below.

Results

Peak functions from single-cue and compound trials in the T_SL_L and L_ST_L groups are plotted in Fig. 1a–d, respectively. Three potential timing effects can be noted from inspection of the figures. First, when rats in both groups responded only on the long duration's nosepoke during compound trials (i.e., "compound-long-only" responding), responses appear to be shifted leftward relative to trials where the long cue was presented in isolation (i.e., "long-alone" responding). Second, when rats in the T_SL_L group responded only on the short duration's nosepoke during compound trials (i.e., "compound-short-only" responding),

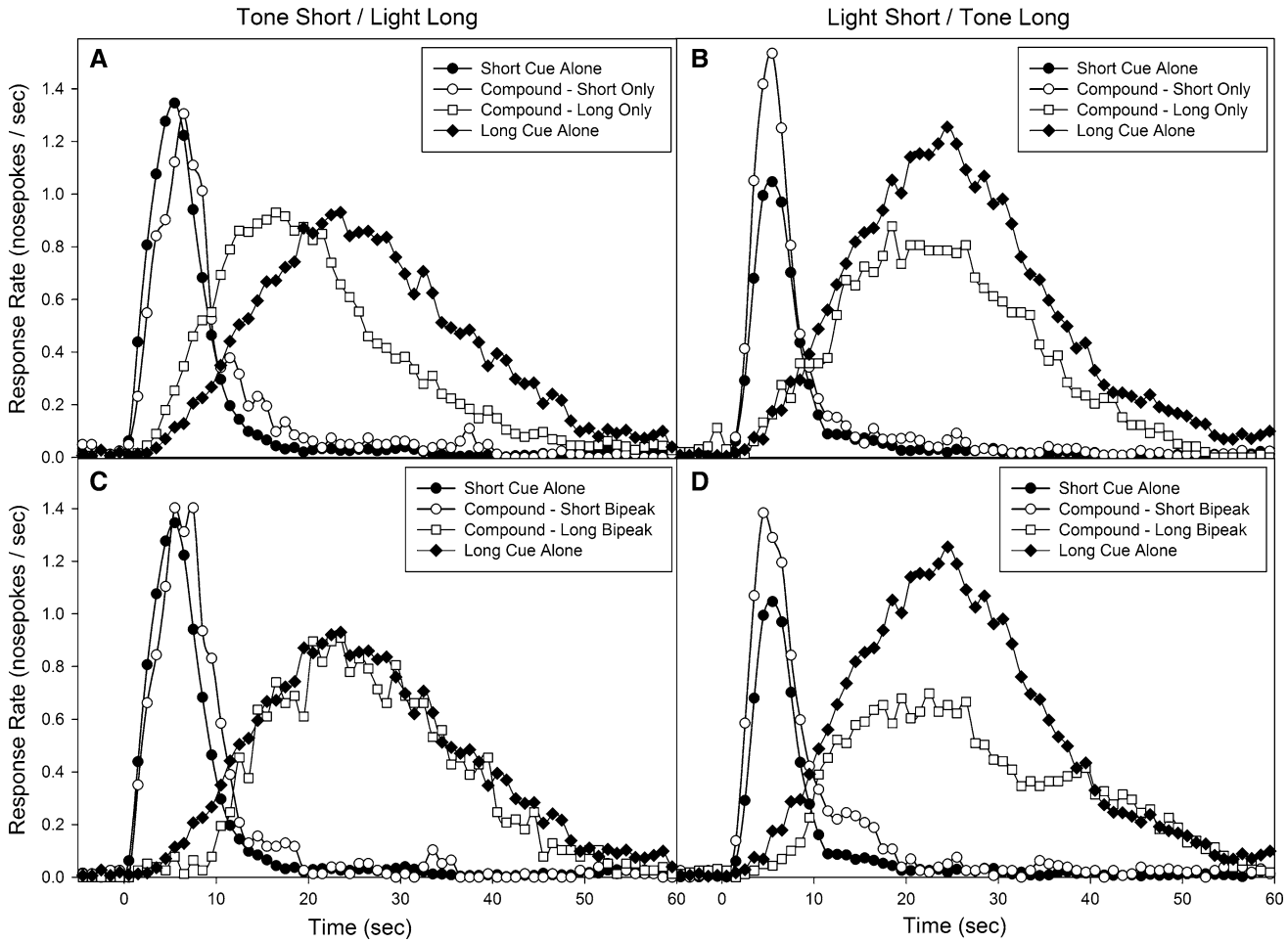

Fig. 1 Peak functions for different trial types in the T_SL_L (**a**, **c**) and L_ST_L (**b**, **d**) groups. Short- and long-alone refer to trials in which the short and long cues were presented in isolation, respectively. *Top panels* plot compound-short-only and compound-long-only responding. *Bottom panels* show compound-short-bipeak and compound-long-bipeak responding

responses appear to be shifted rightward relative to trials where the short cue was presented in isolation (i.e., "short-alone" responding). However, this shift appears to be absent in the L_ST_L group. Third, the leftward shift on the long duration's nosepoke appears to be attenuated or absent when rats first responded on the short cue's nosepoke and subsequently switched to responding on the long cue's nosepoke during compound trials (i.e., "compound-short-bipeak" and "compound-long-bipeak" responding).

To evaluate these observations, we first analyzed the peak time data from both groups, which are plotted in Fig. 2. Analyses revealed an expected main effect of Duration, [$F(1, 16) = 713.39$, $p < 0.001$], and a main effect of Compounding [$F(2, 32) = 6.64$, $p < 0.005$]. However, these effects were significantly moderated by Group [Duration × Group, $F(1, 16) = 5.10$] and each other [Duration × Compounding, $F(2, 32) = 15.49$]. The main effect of Nosepoke Pairing was also significant [$F(1, 16) = 10.06$, $p < 0.01$], as responding was earlier when the short duration was associated with the left nosepoke than when it was associated with the right nosepoke (data not shown). This effect did not interact with any other variables. Importantly, the three-way Duration × Group × Compounding interaction was significant [$F(2, 32) = 4.30$, $p < 0.05$]. Therefore, the data from each group were analyzed separately.

In the T_SL_L group, pairwise comparisons confirmed that compound-short-only peak times were later than short-alone peak times ($p < 0.005$). Similarly, short-bipeak responding was marginally later than short-alone responding ($p = 0.058$). Indeed, there was no difference in peak times during compound-short-only and short-bipeak trials. Conversely, compound-long-only peak times were earlier than long-alone responding ($p < 0.001$), whereas long-bipeak times did not differ from long-alone responding. An anonymous reviewer suggested that the shift in compound-long-only responding might reflect an improvement in accuracy for timing during compound trials, as rats in this group appeared to peak later than the objective, 20-s criterion time during long-alone

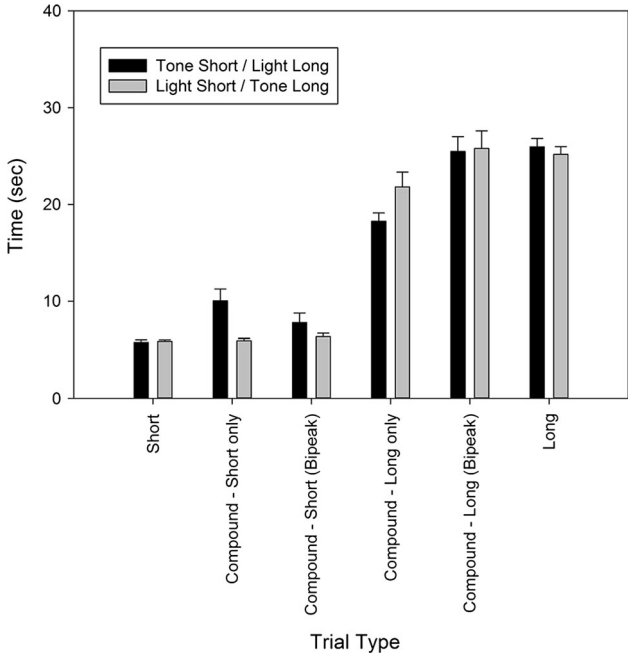

Fig. 2 Peak times for different trial types, split by group. Trial types are defined as in Fig. 1. *Error bars* reflect standard errors

trials. However, this is doubtful for several reasons. First, a one-sample t test on compound-long-only trials showed a trend for peak times to be shifted leftward from the 20-s criterion time ($p = 0.084$, with 8/10 rats having peak times on these trials earlier than 20 s), suggesting that rats were not timing the long duration veridically during these trials. In contrast, long-bipeak peak times were not significantly different from long-alone peak times and were rightward shifted from 20 s ($p < 0.01$). Furthermore, both compound-short-only and compound-bipeak peak times were rightward shifted from the 5-s criterion duration ($p < 0.001$ and $p < 0.05$, respectively). Taken together, these data strongly suggest that rats were not timing accurately during compound trials.

In the L_ST_L group, neither compound-short-only, nor short-bipeak, peak times differed from short-alone peak times. However, compound-long-only peak times were significantly earlier than long-alone peak times ($p < 0.05$). Long-bipeak responding did not differ from long-alone peak times. One-sample t tests indicated that all responding on the short nosepoke was later than the criterion time of 5 s, as was long-alone and long-bipeak responding (all $ps < 0.05$). There was no significant difference between long-only peak times and the programmed criterion duration of 20 s.

Single-trials analysis

To further evaluate temporally controlled behavior during compound trials, we next analyzed data obtained from the single-trials analyses. Start times for the two groups are plotted as a function of trial type in Fig. 3. Start times in the T_SL_L group appear to show similar trends to the peak time data described above. Specifically, compound responding on the short nosepoke appears to be shifted rightward relative to trials where the short cue was presented in isolation. Furthermore, compound responding on the long nosepoke appears to be shifted leftward relative to trials where the long cue was presented alone. However, this effect only appears to emerge when rats responded solely on the long duration's nosepoke during compound trials (i.e., during compound-long-only trials). In contrast to these potential effects, start times in the L_ST_L group appear to have been unaffected by the presence of the compound cue.

These impressions were confirmed by statistical analysis. The expected main effects of Duration, [$F(1, 16) = 334.32$, $p < 0.001$], and Compounding, [$F(2, 32) = 10.04$, $p < 0.001$], were significant. In addition, significant Duration × Group, [$F(1, 16) = 9.4$, $p < 0.01$], and Duration × Compounding, [$F(2, 32) = 14.41$, $p < 0.001$], interactions were found. There was also a main effect of Nosepoke Pairing, with rats in the short-left/long-right group starting earlier than rats in the counterbalanced condition, [$F(1, 16) = 8.94$, $p < 0.05$]. Importantly, the three way Duration × Group × Compounding interaction was also significant, [$F(2, 32) = 5.84$, $p < 0.01$]. Therefore, data from each group were analyzed separately.

In the T_SL_L group, pairwise comparisons revealed both compound-short-only and short-bipeak start times were significantly later in comparison with trials where the short

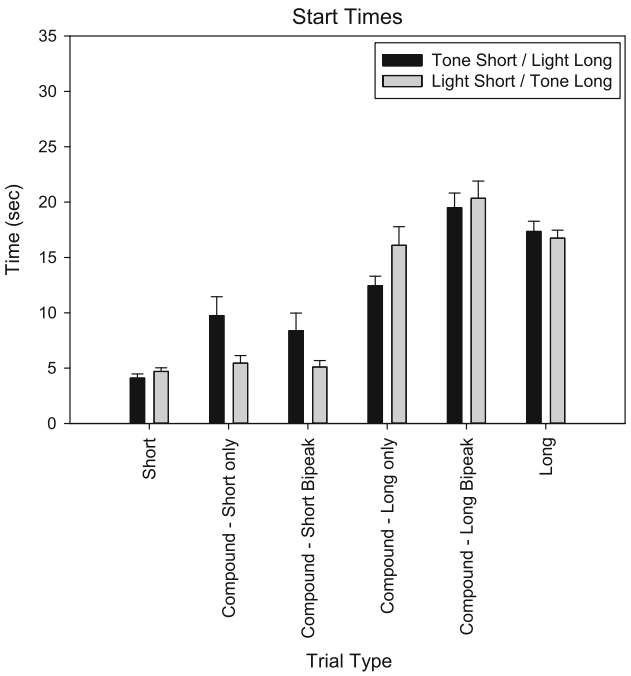

Fig. 3 Start times for different trial types, split by group

cue was presented in isolation ($ps < 0.05$). Start times for these two compound trial types did not differ. Furthermore, compound-long-only start times were significantly earlier relative to both long-bipeak and long-alone trials ($ps < 0.005$). However, long-bipeak and long-alone start times did not differ. No significant differences were found in the L_ST_L group.

Stop times for the two groups are plotted as a function of trial type in Fig. 4. Compound responding in the T_SL_L group appears to mirror the patterns seen for both peak and start times. Specifically, all forms of compound responding on the short duration's nosepoke appear to be shifted rightward relative to trials where the short cue was presented in isolation. Furthermore, compound-long-only responding appears to be leftward shifted when compared to trials in which only the long cue was presented. In the L_ST_L group, short compound responding again appears to be unaffected by the presence of the compound cue, regardless of compound response type. However, like peak times, when rats responded only on the long nosepoke during compound trials, stop times appear to be earlier than those when the long cue was presented in isolation.

These impressions were confirmed by statistical analysis. The main effect of Duration, [$F(1, 16) = 461.8$, $p < 0.001$], was significant, and this variable interacted with Group, [$F(1, 16) = 4.75$, $p < 0.05$]. Furthermore, there was a main effect of Compounding, [$F(2, 32) = 6.94$, $p < 0.005$], that interacted with Duration, [$F(2, 32) = 46.67$, $p < 0.001$], and also Nosepoke Pairing, [$F(2, 32) = 4.8$, $p < 0.05$]. Critically, the three-way Duration × Group × Compounding interaction was also significant [$F(2, 32) = 6.37$, $p < 0.01$]. Therefore, data from each group were analyzed separately.

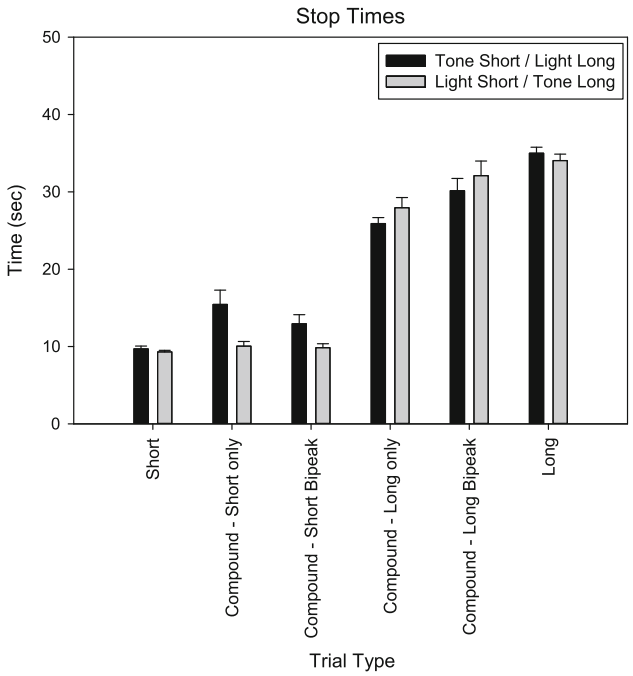

Fig. 4 Stop times for different trial types, split by group

In the T_SL_L group, pairwise comparisons revealed both compound-short-only and short-bipeak stop times were significantly later in comparison with trials where the short cue was presented in isolation ($ps < 0.05$). There was a trend for this shift to be more dramatic during compound-short-only trials than compound-short-bipeak trials ($p = 0.053$). In contrast, compound-long-only and compound-long-bipeak stop times were significantly earlier relative to stop times when the long cue was presented alone ($ps < 0.05$). Furthermore, this shift was more dramatic during compound-long-only trials than compound-long-bipeak trials ($p < 0.05$). In the L_ST_L group, no differences were observed between short-alone and short compound responding. However, compound-long-only stop times were leftward shifted relative to long-alone and compound-long-bipeak stop times ($ps < 0.05$). Compound-long-bipeak and long-alone stop times did not differ.

Examination of scalar variance

Coefficients of variation (CVs) were analyzed next, in order to test for timescale invariance. There was no main effect of Duration or Compounding [$Fs < 1$]. However, Duration × Compounding, [$F(2, 32) = 5.97$, $p < 0.05$], and Compounding × Group, [$F(2, 32) = 10.38$, $p < 0.05$], interactions were found. In addition, a Group × Nosepoke Pairing interaction was seen [$F(1, 16) = 7.51$, $p < 0.05$]. This interaction resulted from narrower peak functions for short-left/long-right rats compared with short-right/long-left rats; however, this only occurred in the L_ST_L group. A three-way Duration × Compounding × Group interaction was found [$F(2, 32) = 3.92$, $p < 0.05$]. Therefore, data from each group were analyzed separately.

In the T_SL_L group, the short-alone peak function was broader than the compound-short and bipeak functions (both short and long) ($ps < 0.05$). However, the long-alone CV did not differ from any trial type. In L_ST_L rats, short-alone peak functions were narrower than short-bipeak functions ($p < 0.05$). No other significant differences were observed.

Response type analysis

We next analyzed whether rats showed a preference for responding on one nosepoke over another during compound trials. To reiterate, we found that the vast majority (98 %) of response patterns during compound trials could be grouped into three categories: short-only, long-only, and bipeak. "Short-only" and "long-only" refer to compound

trials where rats responded solely on the short or long nosepokes, respectively. "Bipeak" refers to compound trials where the rats began responding on the nosepoke associated with the short cue before switching to the nosepoke associated with the long cue.

The proportions of trials in which rats engaged in each pattern of responding are plotted in Fig. 5. Visual inspection of the graph suggests that short-only responses were the most common response type in the L_ST_L group, whereas long-only responses were the most common response pattern in the T_SL_L group. In other words, rats in both groups showed a strong tendency to respond solely on the nosepoke associated with the houselight. In addition, there appears to be a higher proportion of bipeak responding in the L_ST_L than T_SL_L group.

These observations were confirmed by statistical analysis. There was a main effect of Group, $F(1, 16) = 39.4$, $p < 0.001$, and this was significant for both $B1$ (i.e., IRL balance comparing both short and long-only responding to bipeak responding), [$F(1, 16) = 14.99$, $p < 0.001$], and $B2$ (i.e., IRL balance comparing short-only responding to long-only responding), [$F(1, 16) = 82.61$, $p < 0.001$]. There was also a main effect of Nosepoke Pairing, [$F(1, 16) = 6.14$, $p < 0.05$], which held across $B1$, [$F(1, 16) = 7.7$, $p < 0.05$], and $B2$, [$F(1, 16) = 9.3$, $p < 0.05$]. However, these latter effects were negligible in extent (e.g., 4 % fewer bipeaks when the 5- and 20-s durations were associated with the left and right nosepokes, respectively). Finally, no Group × Nosepoke Pairing interaction was found, [$F = 3.5$].

Response rate analysis

Compounding also appeared to have exerted effects on response rate that differed as a function of group (see Fig. 1). Specifically, in the L_sT_L group, compound responding on the short nosepoke appears to be elevated relative to when the short cue was presented alone. In contrast, response rates on the long nosepoke appear to be lower during compound trials relative to when the long cue was presented in isolation. Both of these effects appear to be absent in the T_sL_L group.

These impressions were confirmed by statistical analysis. There was a significant main effect of Duration [$F(1, 16) = 15.66$, $p < 0.005$], as response rates were higher on the short duration's nosepoke than the long. There was also a main effect of Nosepoke Pairing, with response rates being higher in the short-left/long-right group, [$F(1, 16) = 8.94$, $p < 0.05$]. In addition, there was a three-way Duration × Compounding × Nosepoke Pairing interaction with responding during long-bipeak trials being lower relative to both long-alone and long-only trials in the short-left/long-right group [$F(2, 32) = 4.3$, $p < 0.05$]. Finally, the critical three-way Duration × Group × Compounding interaction was also significant, [$F(2, 32) = 7.59$, $p < 0.005$]. Therefore, each group was analyzed separately.

Effects in the T_sL_L group primarily reflected the main effect of Duration. Both short-alone and short-bipeak response rates were higher than all patterns of responding on the long nosepoke (i.e., long-alone, long-only, and long-bipeak; ps < 0.05). Similarly, short-only response rates were higher than long-alone responding ($p < 0.05$).

Similar effects were observed in the L_sT_L group. For example, short-only and short-bipeak response rates were higher relative to long-bipeak trials (ps < 0.05), and there was a trend for short-alone responding to be higher than long-bipeak responding ($p = 0.06$). However, unlike the T_sL_L group, there were also within-duration differences. Specifically, short-only responding was higher relative to trials where the short cue was presented in isolation ($p < 0.05$), and there was a trend for this effect to be present for short-bipeak trials as well ($p = 0.06$). Conversely, long-bipeak response rates were lower than long-alone and long-only responding ($p < 0.05$). However, the apparent difference between long-alone and long-only responding was not significant ($p > 0.05$).

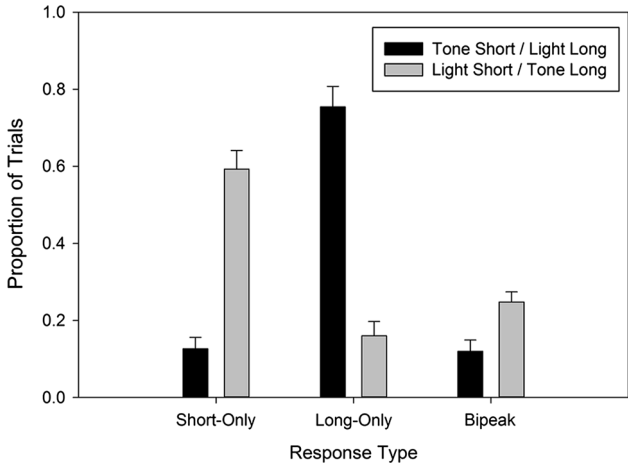

Fig. 5 Proportion of trials of each trial type, split by group

Discussion

This study was intended to test predictions made by the averaging and simultaneous temporal processing (STP) hypotheses regarding how compound responding should proceed when each component cue is associated with a distinct manipulandum and duration. According to the averaging hypothesis, when presented with a compound cue, rats compute a probability-weighted average of the anchor durations (Swanton and Matell 2011; Matell and

Kurti 2014). If this occurred in the current study, responding on either aperture should have been guided by the same, average temporal expectation. Therefore, compound response distributions at the short and long duration's nosepokes should have overlapped.

In contrast, according to the STP hypothesis, when presented with a compound cue, rats time the anchor durations simultaneously and veridically. If this hypothesis was correct, rats in the current study should have progressed from responding on the nosepoke associated with the short duration to the nosepoke associated with the long duration during compound trials. Furthermore, response distributions at each nosepoke should have been centered over their respective criterion times.

Some limited support for the STP hypothesis was found in the current data. Specifically, rats occasionally responded on both nosepokes in sequence during compound trials, and this led to distinct peak functions at either aperture with peak times that were often centered over their respective criterion times. However, in the T_SL_L group, temporally controlled responding during such "bipeak" trials was not entirely veridical, as responses on the short nosepoke were shifted toward the long duration. Furthermore, bipeak responding was only seen on a small proportion of compound trials (12 and 25 % in the T_SL_L and L_ST_L groups, respectively).

Instead, on the majority of compound trials (88 and 75 % for the T_SL_L and L_ST_L groups, respectively), rats responded solely on the short or long duration's nosepoke. During these trials, responses on a given aperture were typically biased toward the duration associated with the other nosepoke. This was seen for both durations in the T_SL_L group, but only for the long duration in the L_ST_L group. Specifically, when rats in the T_SL_L group responded solely on the short duration's nosepoke during compound trials, responding was rightward shifted relative to trials where the short cue was presented alone (median peak time shift = 56 %). Conversely, when these rats only responded on the long nosepoke during compound trials, peak times were leftward shifted compared with long-cue trials (median shift = 34 %). Similarly, in the L_ST_L group, when rats responded solely on the long nosepoke during compound trials, peak times were leftward shifted relative to long-cue trials (median shift = 18 %). However, when these rats responded solely on the short nosepoke during compound trials, peak times were not shifted (median shift = −3 %).

When present, shifts in peak times for each nosepoke were in the direction of the duration associated with the other nosepoke. Therefore, these responses reflected a combination of both durations, which is consistent with temporal averaging behavior. However, rats did not appear to be timing the same reinforcement probability-weighted average when responding on either aperture, as predicted by the averaging hypothesis. Rather, as the peak time on each nosepoke fell closer to its respective duration, more weight appeared to be given to the duration associated with the nosepoke on which rats responded. We will refer to this as "selection-dependent" weighting, as rats appeared to select one of the two cues and weighted its duration more heavily.

To summarize, during the majority of compound trials, rats responded exclusively on one or the other nosepoke at a time that appeared to be a synthesis of both durations, with more weight being given to the duration associated with the nosepoke on which the rat responded. In other words, rats over-produced the short duration and under-produced the long duration. Similar effects have been documented in the past. For example, when Parkinson's disease patients are trained to time two different intervals on different trials while on their dopaminergic medication and later tested off-medication, a "migration effect" occurs, whereby the shorter and longer durations are over- and under-produced, respectively (Malapani et al. 1998, 2002). Furthermore, a related and well-known effect referred to as "Vierordt's law" suggests that these patterns are even seen in healthy subjects (Vierordt 1868), albeit to a lesser degree. Specifically, when subjects are asked to make duration judgments regarding multiple intervals across different trials, shorter and longer durations tend to be under- and over-estimated, respectively (Bausenhart et al. 2014; Mayer et al. 2014). This pattern has been observed across a variety of timing tasks and sensory modalities (for discussion see Lejeune and Wearden 2009). Subjects exhibiting the migration effect or Vierordt's law appear to time a weighted combination of multiple durations, with more weight being given to the duration they intend to time. Therefore, these effects share similarities with the selection-dependent weighting seen in the current data.

Our results also provide insight into factors that might impact this selection process. To reiterate, during compound trials in the present study, rats responded in one of three ways: solely on the short duration's nosepoke, solely on the long duration's nosepoke, or on both nosepokes in sequence. On the majority of these trials (81 %), rats responded on only one of the two nosepokes. Of these "single-nosepoke" trials, rats in the T_SL_L group responded on the long nosepoke 85 % of the time. In contrast, rats in the L_ST_L group responded on the short nosepoke 78 % of the time. Therefore, rats in both groups appeared to primarily attend to (i.e., "select") the duration associated with the light during compound trials.

Why would this be the case? One possibility is that the light was a more salient stimulus than the tone and, thus, dominated responding, due to overshadowing. However, pilot studies from our laboratory have shown that

manipulating the relative salience of the two cues has no impact on compound responding, which casts doubt on this explanation. An alternative explanation is that rats implicitly associated the light more strongly with reward than the tone. This is consistent with past stimulus compounding literature showing that rats are more likely to associate positive outcomes with visual than auditory stimuli. For example, using an operant paradigm, Weiss et al. (1993) trained rats that a tone–light compound cue was associated with either a hedonically positive or negative outcome. Test trials were also included in which either cue was presented in isolation. During test trials, rats responded more to the light when the compound cue had been associated with a positive outcome. In contrast, when the compound outcome was negative, responses were higher to the tone. As both cues in the current study were associated with positive outcomes (reward), this finding may account for the light bias seen in our data.

The potentially heightened value of visual stimuli relative to auditory cues may also help explain why rats in the L_ST_L group appeared to time the short duration veridically during compound trials. Specifically, we propose that different response strategies become active during compound trials depending on both the temporal reliability and implicit value of each of the two cues. For example, in the L_ST_L group, the short cue was presumably both highly valued, given that it was a visual stimulus, and highly reliable, given that shorter durations are timed with less absolute error due to scalar variance. The conjunction of these two features may have caused rats in this group to engage in what we will refer to as an "exclusive strategy" on the majority of compound trials. This entails giving full weight to a selected duration (i.e., timing the duration veridically). Presumably, this strategy is activated during compound trials when one cue is perceived as a far more profitable source of temporal information relative to the other.

However, the short cue in the L_ST_L group is the only case where these two "desirable" features occurred together (i.e., high reliability and high value). For example, in the T_SL_L group, the short cue presumably had higher reliability relative to the long duration, yet a lower value, given that it was an auditory stimulus. In contrast, the long cue had lower reliability and higher value, given that it was a visual stimulus. Finally, in the L_ST_L group, the long cue (i.e., the tone) presumably had both low reliability and low value.

When rats timed a duration associated with one or more of these "undesirable" features during compound trials, they appeared to engage in what we will refer to as a "biasing strategy" such that some weight is given to the other, non-timed duration. Here, while the selected duration is weighted heavily, it is not weighted fully as in the exclusive strategy, causing responses to be shifted toward the other duration. Presumably, biasing emerges when there is uncertainty regarding the likelihood that timing a given cue will result in reward.

Responding during bipeak trials, during which rats responded first on the short nosepoke before switching to the long nosepoke, provides some support for this latter point. Specifically, during bipeak trials, timed responding on the short nosepoke was similar to compound trials where rats responded solely on this manipulandum, with T_SL_L and L_ST_L rats using the biasing and exclusive strategies, respectively. However, bipeak responding on the long nosepoke was veridical in both groups, which contrasts to the biasing seen when rats responded solely on this aperture during compound trials. If biasing emerges when there is uncertainty regarding which duration to time during a trial, this result might be expected, as responding at the short nosepoke would have allowed rats to evaluate whether the short duration was informative in regard to when reward would be delivered. When responses on this nosepoke failed to produce reward, rats may have reoriented toward timing the long duration without bias, as any uncertainty regarding whether the short duration should be taken into account would have been eliminated.

Interestingly, the conjunction of temporal reliability and intrinsic value of a given cue also appeared to correlate with the response rate observed on its associated nosepoke during compound trials. For example, in the L_ST_L group, responding on the short duration's nosepoke (i.e., the response associated with the light) was elevated relative to trials where the short cue was presented in isolation, and correspondingly, the light was presumably associated with both high reliability and high value. Similarly, the tone was associated with both low value and low reliability, and compound responding on its associated nosepoke was lower relative to trials where the tone was presented alone. Furthermore, in the T_SL_L group, both cues were associated with one desirable and one undesirable feature, and responding on each nosepoke was roughly equivalent to cue-alone responding. It is plausible that differences in the conjunction of reliability and value caused these response rates differences, in addition to impacting strategy selection. However, admittedly, this explanation assumes that these same factors did not impact response rates to the same degree when each cue was presented in isolation. Why this would be the case is unclear.

One possibility relates to the "contrast effect," in which responding for an outcome depends on the value of other outcomes that are concurrently present or are anticipated to become available in the near future. For example, in Wright et al. (2013), rats were given access to either a 4 % sucrose solution for two minutes followed by a 32 % solution for another two minutes during some sessions or a

4 % sucrose solution for four minutes during other sessions. Each session type was signaled by distinct contextual cues. Rats consumed less of the 4 % solution and responded in shorter bouts when it was followed by the more preferred 32 % solution relative to sessions where only the 4 % solution was available. In other words, the expectation of the 32 % solution appeared to devalue the 4 % solution.

A contrast effect could have also occurred during compound trials in our study, wherein the response rate on a particular nosepoke was determined by evaluating the properties of its associated cue relative to those of the other stimulus. For example, during compound trials in the L_ST_L group, the contrast between the proposed high value and reliability of the light and the low value and reliability of the tone may have resulted in heightened response rates on the light's nosepoke relative to trials where the light was presented in isolation. This same contrast would also account for the lower response rates observed in the tone's nosepoke during compound trials. Furthermore, in the T_SL_L group, no contrast would be expected, as both cues possessed one desirable and undesirable feature, and accordingly, compound response rates on either nosepoke were similar to cue-alone trials.

Finally, we note that several prior compounding findings from our laboratory, in which a single nosepoke has been used, can be explained by assuming rats engage in the same response strategies during compound trials as the ones observed here. For example, in all previous studies, rats in T_SL_L groups have produced scalar compound response distributions falling in between the two criterion durations, but biased toward the long light duration. However, this pattern of temporal averaging is not usually observed in L_ST_L groups (Swanton and Matell 2011; Kurti et al. 2013; Matell and Kurti 2014). Instead, the left half of these rats' compound peak functions, as well as the peak time, typically overlaps the short-anchor distribution (i.e., the peak function when the short cue is presented alone). Furthermore, the right half of the compound peak function is often elongated and typically falls in between the right tails of the short and long anchor distributions. Consequently, these peak functions are positively skewed and non-scalar.

One potential explanation for these results is that, on the majority of compound trials, L_ST_L rats time the light in a veridical manner (i.e., used the exclusive strategy), similar to the L_ST_L group of the present study. However, on a small proportion of trials, rats may have timed the long duration in a biased manner (i.e., used the biasing strategy). When the ratio of short to long durations is small (e.g., 1:3), the response functions will partially overlap. Furthermore, because of the predominance of the exclusive strategy, averaging over trials would result in a peak function in which the left half overlaps the short-anchor peak function, whereas the right half falls in between the anchor distributions. Consistent with this interpretation, when Swanton and Matell (2011) used a large ratio between the short and long durations (1:6), L_ST_L rats' compound peak functions became bimodal, with one peak centered over the short criterion time and a second falling in between the two durations.

Another finding that can be accounted for by these strategies comes from Matell and Kurti (2014) who evaluated the effects of manipulating cue value on temporal averaging. Specifically, in Experiment 2, the authors used T_SL_L groups (i.e., rats who usually display scalar averaging) and manipulated the value of each cue by systematically altering the reinforcement density of each stimulus. When the reinforcement densities of the two cues were equal, the rats showed scalar temporal averaging. In contrast, when the short cue had a higher reinforcement density than the long cue, compound response distributions became positively skewed and non-scalar. Critically, when the reinforcement densities were highly discrepant, the left half of the compound peak function overlapped that of the short-anchor duration, and the right half fell in between the two criterion times. Thus, their results mirrored those seen in L_ST_L groups at smaller ratios. The authors concluded that, when the tone was perceived as more valuable than the light, rats primarily timed the tone in a veridical manner, rather than engaging in averaging behavior. This is equivalent to the exclusive strategy we propose here and supports the notion that the use of different strategies is affected by cue value. As L_ST_L groups show this pattern of responding even when the reinforcement densities for the two cues are equal, the authors suggested that L_ST_L rats might naturally perceive the light as more valuable than the tone during compound trials, which is also consistent with the proposal we offer here.

We cannot definitively state that subjects use the same compound response mechanisms when single or multiple response options are provided based on the current data alone. However, the parallels between compound responding observed here and in past single-nosepoke studies suggest this may be the case. Furthermore, this assumption is more parsimonious than assuming that different timing mechanisms are invoked depending on the number of response options employed in a given study. Indeed, the use of different procedures for studying timing in animals (e.g., bisection procedure and peak procedure) implicitly assumes that the same temporal mechanisms are utilized, regardless of task demands.

In conclusion, the current results suggest that patterns of compound responding seen in the past have not been solely due to rats timing two durations veridically, as suggested by the STP hypothesis. Rather, on the majority of trials, rats appear to engage in temporal averaging behavior. However, the average duration being timed did not appear

to be solely determined by the reinforcement probabilities associated with either cue. Therefore, our results provide novel insight into the mechanisms by which rats weight different durations when integrating. Broadly, these results suggest that this weighting process is more flexible than previously conceived.

Such flexibility is consistent with averaging behavior seen in other domains. For example, pigeons presented with multiple landmarks that provide discrepant information regarding the location of a desired goal will compute a vector average of each landmark's distance and direction toward its associated goal location (Cheng 1988, 1989). Importantly, landmarks that have been closer to the desired goal in the past will be weighted more heavily than farther landmarks, presumably because closer landmarks provide more reliable spatial information. Similarly, desert ants show reliability-based weighting while navigating when celestial and terrestrial spatial cues are placed in conflict with one another (Legge et al. 2014). Furthermore, humans have been shown to integrate discrepant spatial (Battaglia et al. 2003), spatiotemporal (Cheng et al. 1996), size (Ernst and Banks 2002; Helbig and Ernst 2007), shape (Helbig et al. 2012), and depth (Jacobs 1999) information. These results are often interpreted within the context of Bayesian decision theory, wherein the variance in the information provided by a given source determines how much weight it is given, with less variable sources being weighted more heavily. Given these results, an important future direction for timing research will be to evaluate factors that influence the combinatorial strategies used when discrepant temporal information is present within the environment.

Compliance with ethical standards

Ethical standards All research presented herein complied with current laws regarding animal welfare in the United States of American and regulations implemented by Villanova University's Institutional Animal Care and Use Committee.

Conflict of interest The authors declare no conflicts of interest.

References

Aitchison J (1982) The statistical analysis of compositional data. J R Stat Soc Series B Stat Methodol 44:139–177

Aitchison J (2005) Compositional data analysis: where are we and where should we be heading? Math Geol 37:829–850

Balsam P, Gallistel C (2009) Temporal maps and informativeness in associative learning. Trends Neurosci 32:73–78

Battaglia PW, Jacobs RA, Aslin RN (2003) Bayesian integration of visual and auditory signals for spatial localization. J Opt Soc Am A 20:1391–1397

Bausenhart KM, Dyjas O, Ulrich R (2014) Temporal reproductions are influenced by an internal reference: explaining the Vierordt effect. Acta Psychol 147:60–67

Cheng K (1988) Some psychophysics of the pigeon's use of landmarks. J Comp Physiol A 162:815–826

Cheng K (1989) The vector sum model of pigeon landmark use. J Exp Psychol Anim Behav Process 15:366–375

Cheng K, Spetch ML, Miceli P (1996) Averaging temporal duration and spatial position. J Exp Psychol Anim Behav Process 22:175–182

Church RM, Meck WH, Gibbon J (1994) Application of scalar timing theory to individual trials. J Exp Psychol Anim Behav Process 20:135–155

Egozcue JJ, Pawlowsky-Glahn V (2005) Groups of parts and their balances in compositional data analysis. Math Geol 37:795–828

Engle MA, Rowan EL (2013) Interpretation of Na–Cl–Br systematics in sedimentary basin brines: comparison of concentration, element ratio, and isometric log-ratio approaches. Math Geosci 45:87–101

Ernst MO, Banks MS (2002) Humans integrate visual and haptic information in a statistically optimal fashion. Nature 415:429–433

Gallistel C, King A, McDonald R (2004) Sources of variability and systematic error in mouse timing behavior. J Exp Psychol Anim Behav Process 30:3–16

Gibbon J (1977) Scalar expectancy theory and Weber's law in animal timing. Psychol Rev 84:279–325

Gibbon J, Balsam P (1981) Spreading association in time. In: Locurto C, Terrace H (eds) Autoshaping and conditioning Theory. Academic Press, London, pp 219–253

Gibbon J, Church RM (1990) Representation of time. Cognit 37:23–54

Gibbon J, Church RM, Meck WH (1984) Scalar timing in memory. Ann N Y Acad Sci 423:52–77

Harris JA (2015) Changes in the distribution of response rates across the CS-US interval: evidence that responding switches between two distinct states. J Exp Psychol Anim Learn Cogn 41:217–231

Helbig HB, Ernst MO (2007) Optimal integration of shape information from vision and touch. Exp Brain Res 179:595–606

Helbig HB, Ernst MO, Ricciardi E, Pietrini P, Thielscher A, Mayer KM, Shultz J, Noppeney U (2012) The neural mechanisms of reliability weighted integration of shape information from vision and touch. NeuroImage 60:1063–1072

Jacobs RA (1999) Optimal integration of texture and motion cues to depth. Vision Res 39:3621–3629

Kacelnik A, Brunner D, Gibbon J (1990) Timing mechanisms in optimal foraging: some applications of scalar expectancy theory. In: Hughes R (ed) Behavioural mechanisms of food selection. Springer, Berlin, pp 61–82

Kurti A, Swanton D, Matell M (2013) The potential link between temporal averaging and drug-taking behavior. In: Arstila V, Lloyd D (eds) Subjective time: the philosophy, psychology, and neuroscience of temporality. MIT Press, Cambridge, pp 599–620

Leak TM, Gibbon J (1995) Simultaneous timing of multiple intervals: implications of the scalar property. J Exp Psychol Anim Behav Process 21:3–19

Legge ELG, Wystrach A, Spetch ML, Cheng K (2014) Combining sky and earth: desert ants (*Melophorus bagoti*) show weighted integration of celestial and terrestrial cues. J Exp Biol 217:4159–4166

Lejeune H, Wearden JH (2009) Vierordt's the experimental study of the time sense (1868) and its legacy. Eur J Cogn Psychol 21:941–960

Macdonald CJ, Cheng RK, Meck WH (2012) Acquisition of "start" and "stop" response thresholds in peak-interval timing is differentially sensitive to protein synthesis inhibition in the dorsal and ventral striatum. Front Integr Neurosci 6:1–16

Malapani C, Rakitin B, Levy R, Meck WH, Deweer B, Dubois B, Gibbon J (1998) Coupled temporal memories in Parkinson's disease: a dopamine-related dysfunction. J Cogn Neurosci 10:316–331

Malapani C, Deweer B, Gibbon J (2002) Separating storage from retrieval dysfunction of temporal memory in Parkinson's disease. J Cogn Neurosci 14:311–322

Martín-Fernández JA, Barceló-Vidal C, Pawlowsky-Glahn V (2003) Dealing with zeros and missing values in compositional data sets using nonparametric imputation. Math Geol 35:253–278

Matell MS, Kurti AN (2014) Reinforcement probability modulates temporal memory selection and integration processes. Acta Psychol 147:80–91

Matell MS, Bateson M, Meck WH (2006) Single-trials analyses demonstrate that increases in clock speed contribute to the methamphetamine-induced horizontal shifts in peak-interval timing functions. Psychopharmacology 188:201–212

Mayer KM, Di Luca M, Ernst MO (2014) Duration perception in crossmodally defined intervals. Acta Psychol 147:2–9

Meck WH, Church RM (1984) Simultaneous temporal processing. J Exp Psychol Anim Behav Process 10:1–29

Meck WH, Cheng R, MacDonald CJ, Gainetdinov RR, Caron MG, Çevik MÖ (2012) Gene-dose dependent effects of methamphetamine on interval timing in dopamine-transporter knockout mice. Neuropharmacology 62:1221–1229

Olton DS, Wenk GL, Church RM, Meck WH (1988) Attention and the frontal cortex as examined by simultaneous temporal processing. Neuropsychologia 26:307–318

Pang KCH, Yoder RM, Olton DS (2001) Neurons in the lateral agranular frontal cortex have divided attention correlates in a simultaneous temporal processing task. Neuroscience 103:615–628

Pawlowsky-Glahn V, Egozcue J (2006) Compositional data and their analysis: an introduction. Geol Soc Spec Publ 264:1–10

Roberts S (1981) Isolation of an internal clock. J Exp Psychol Anim Behav Process 7:242–368

Swanton DN, Matell MS (2011) Stimulus compounding in interval timing: the modality–duration relationship of the anchor durations results in qualitatively different response patterns to the compound cue. J Exp Psychol Anim Behav Process 37:94–107

Swanton DN, Gooch CM, Matell MS (2009) Averaging of temporal memories by rats. J Exp Psychol Anim Behav Process 35:434–439

Vierordt K (1868) Der Zeitsinn nach Versuchen, Germany

Weiss SJ, Panlilio LV, Schindler CW (1993) Single-incentive selective associations produced solely as a function of compound-stimulus conditioning context. J Exp Psychol Anim Behav Process 19:284–294

Wright RL, Gilmour G, Dwyer DM (2013) Microstructural analysis of negative anticipatory contrast: a reconsideration of the devaluation account. Learn Behav 41:353–359