Psychological Research
DOI 10.1007/s00426-015-0675-8

ORIGINAL ARTICLE

The roles of stimulus and response uncertainty in forced-choice performance: an amendment to Hick/Hyman Law

Tim Wifall[1] · **Eliot Hazeltine**[1] · **J. Toby Mordkoff**[1]

Received: 10 January 2014 / Accepted: 27 May 2015
© Springer-Verlag Berlin Heidelberg 2015

Abstract Hick/Hyman Law describes one of the core phenomena in the study of human information processing: mean response time is a linear function of average uncertainty. In the original work of Hick, (1952) and Hyman, (1953), along with many follow-up studies, uncertainty regarding the stimulus and uncertainty regarding the response were confounded such that the relative importance of these two factors remains mostly unknown. The present work first replicates Hick/Hyman Law with a new set of stimuli and then goes on to separately estimate the roles of stimulus and response uncertainty. The results demonstrate that, for a popular type of task—visual stimuli mapped to vocal responses—response uncertainty accounts for a majority of the effect. The results justify a revised expression of Hick/Hyman Law and place strong constraints on theoretical accounts of the law, as well as models of response selection in general.

Introduction

One important finding to usher in the age of information processing is the phenomenon now known as Hick/Hyman Law (HHL; Hick, 1952; Hyman, 1953), which holds that mean response time (RT) is a linear function of average uncertainty (H), regardless of how uncertainty is manipulated:

$$\text{mean RT} = a + bH, \tag{1}$$

where mean RT is for the block of trials, H is average uncertainty for the block (in bits), and a and b are the regression constants estimated from the data. The calculation of H derives from Information Theory (Shannon, 1948; Shannon & Weaver, 1949) which, in turn, took the term from statistical mechanics, where H is the symbol for entropy. The general formula for H (when the units are bits) is

$$H = -\sum p(i) \log_2 p(i), \tag{2}$$

where $p(i)$ refers to the probability of each of the possible types of trial. The quantity $-\log_2 p(i)$ is often referred to as the surprisal value for a particular trial; thus, average uncertainty can also be thought of as the weighted mean surprisal value for all of the trials in a block. (In fact, "weighted mean surprisal value" is probably the much better label for H, as it specifies which measure of central tendency should be used, but, unfortunately, the label "average uncertainty" is what has been used in the literature.) In the case where all possible trials are equally frequent (as in the present experiments), the general formula for average uncertainty (in bits) reduces to

$$H = \log_2 m, \tag{3}$$

where m is the number of different types of trial. This is why HHL is often expressed in a way that conceals the central role of the concept of average uncertainty:

$$\text{mean RT} = a + b \log_2 m. \tag{4}$$

There has been previous work that has established both the widespread generality of HHL as well as some limits (see, e.g., Attneave, 1959; Kveraga, Boucher, & Hughes, 2002; Leonard, 1959; Welford, 1968). For example, when stimuli and responses have a high degree of compatibility

---

Tim Wifall is currently working at Hitachi America Ltd., R&D.

✉ Tim Wifall
twifall@gmail.com

[1] Department of Psychology, University of Iowa, E11 Seashore Hall, Iowa City, IA 52241, USA

Published online: 10 June 2015

(Leonard, 1959; Teichner & Krebs, 1974; Wright, Marino, Belovsky, & Chubb, 2007), including tasks in which participants must make saccades towards a target (Kveraga et al., 2002), the slope, *b*, can be close to or equal to 0. This is not a violation of the law, per se, but does place a limit on its applicability. Somewhat more problematic have been the occasional findings of non-linearity when atypical tasks are employed (e.g., Longstreth, El-Zahhar, & Alcorn, 1985), which imposes another limit on the law and implies that any new task should be tested for linearity before being used. Finally, there is the suggestion that the law ceases to be linear when uncertainty is more than 3 bits (e.g., Longstreth, 1988), which places an upper bound on HHL.

Recently, interest in HHL has been renewed, with a focus on what the phenomenon reveals about underlying perceptual, decisional, and/or response-related mechanisms (e.g., Hawkins, Brown, Steyvers, & Wagenmakers, 2012; Schneider & Anderson, 2011; Usher, Olami, & McClelland, 2002;). The phenomenon of RT increasing as a linear function of the log of the number of S–R alternatives has served as a benchmark for various models of S–R translation, but these models differ in terms of the type of uncertainty that are proposed to drive the law: some posit that increases in the number of S–R mappings (and, therefore, the number of possible stimuli) lead to longer RTs (e.g., Jamieson & Mewhort, 2009; Schneider & Anderson, 2011), whereas others claim that the number of responses is critical (e.g., Usher et al., 2002). Note that in the original Hick and Hyman experiments, each stimulus was associated with a different (unique) response. Thus, the number of stimuli and the number of responses were confounded, so it could not be determined whether the (linear) changes in mean RT were due to changes in uncertainty regarding which stimulus would be presented or to changes in uncertainty as to which response would be required. Both Hick, (1952) and Hyman, (1953) suggested that the changes in mean RT resulted from changes in stimulus uncertainty, but they had no direct evidence to support this claim over the alternative proposal that response uncertainty drives the effect. Given the renewed interest in this general question—especially given how several recent models embody specific claims as to which type of uncertainty should be more important—the present study was designed to directly compare the roles of stimulus and response uncertainty in a popular form of the forced-choice task.

Previous investigations of stimulus and response uncertainty

The question of whether stimulus or response uncertainty plays the larger role in HHL has been examined before (e.g., Keele, 1970; Pollack, 1959; Rabbitt, 1959; see, also, LaBerge & Tweedy, 1964; LaBerge, Legrand, & Hobbie, 1969). For example, Rabbitt, (1959) had participants perform a card-sorting task in which stimulus and response uncertainty were independently manipulated. Stimulus uncertainty was manipulated by changing the total number of types of cards, whereas response uncertainty was manipulated by changing the number of piles into which the cards were to be sorted. The results were somewhat complex, perhaps because of the atypical nature of the task. When there were more than two stimuli per pile, the time to sort was primarily dependent on the number of piles and relatively unaffected by the number of stimuli associated with each pile, suggesting that response uncertainty plays the larger role. However, when the one- and two-stimuli-per-pile conditions were compared, a large difference in sorting time was observed, regardless of the number of piles, suggesting that stimulus uncertainty can be much more important. In summary, the answer to the critical question depended on the specific levels of each of the factors.

There are several aspects of Rabbitt, (1959) study that make its interpretation less than straightforward. First, the non-linear effects that were obtained in the study are inconsistent with the core prediction of HHL. Second, the task was more difficult than most studies of forced-choice performance—requiring card-sorting (and, therefore, aimed hand movements) instead of button-pressing or vocal responses—and, yet, the participants received very little practice. These issues make the results of Rabbitt, (1959) inconclusive with regard to the roles that stimulus and response uncertainty play in more typical and practiced, choice-RT tasks.

An experiment reported by Bricker, (1955), while not explicitly designed to measure the contributions of stimulus and response uncertainty to Hick/Hyman Law, is also relevant. In particular, one of the conditions included by Bricker was argued to involve a manipulation of response uncertainty while holding stimulus uncertainty constant (an approach that is closely matched by the present experiments). Using similar stimuli to those of Hyman, (1953) and identical responses, Bricker mapped the same eight stimulus patterns onto two, four, or eight responses, which allowed for a comparison of the slopes between when stimulus and response uncertainty were both varied (Hyman, 1953, Exp. 1) and when only response uncertainty was varied (Bricker, 1955, Condition C). Because the slopes were almost identical, Bricker concluded that “it appears that reducing stimulus and response uncertainty together decreases reaction time no more than decreasing response uncertainty alone” (ibid, p. 81). In other words, in contrast to the suggestions of both Hick, (1952) and Hyman, (1953), Bricker, (1955) argued that all of Hick/Hyman Law should be associated with changes in response uncertainty.

However, as acknowledged by Bricker, (1955, p. 80), the comparison with Hyman, (1953) may not be appropriate for several reasons. The greatest impediment to comparison comes from the stimuli. While Hyman presented a single element in one of eight different locations, Bricker, (1955) presented three elements (in the condition that was used for the comparison), each in one of two locations, with the eight different "stimuli" being eight different combinations of locations of the three elements. Thus, when Bricker reduced the number of responses, the number of relevant components of the display was also decreased. When all eight responses were involved, the pattern across all three elements needed to be processed. However, when only two responses were involved, only two of the elements needed to be processed in order to determine the correct response. In other words, Bricker, (1955) did not actually hold stimulus uncertainty constant while manipulating response uncertainty, so the comparison between Hyman, (1953, Exp. 1) and Bricker, (1955, Condition C) does not answer our question.

The other studies bearing on this issue have also provided ambiguous results. For example, Pollack, (1959) had participants identify a target word that was embedded in white noise. Before each test, participants were provided with a list of potential target words, ranging from 2 to 32 possibilities. After the presentation of the target word, a card containing the target word along with a varied number of distractor words was placed in front of the participant, who then had to select the target word. Pollack varied the number of response options from 2 to 64. Across all of conditions, the results consistently revealed that message reception was independent of the number of possible messages (i.e., the number of possible stimuli) and dependent only on the number of possible responses. However, this effect was only observed in accuracy, whereas HHL predicts mean RT.

Finally, Keele, (1970) compared RT in conditions with two stimuli mapped to two responses, four stimuli mapped to four responses, and four stimuli mapped to two responses. The four-to-four mapping was significantly slower than the four-to-two mapping, which in turn, was significantly slower than the two-to-two mapping, suggesting that both stimulus and response uncertainty contribute to RT. However, participants again received little practice, and, because there were only two levels of each type of uncertainty, it could not be verified that HHL held under these conditions. In summary, no task that is known to exhibit the basic pattern of HHL (in mean RT) has ever succeeded in separating the contributions of stimulus and response uncertainty.

Overview

The goal of the present research was to isolate the contributions of stimulus and response uncertainty to the basic phenomenon of HHL. To hold stimulus discriminability constant across the levels of the stimulus-uncertainty manipulation, the present work employed a new set of stimuli that was both unfamiliar and arbitrary to the participants. We note that Hick, (1952) and Hyman, (1953) both used spatial arrays of lights, and the stimuli produced from such displays may become more confusable as the number of possible stimuli increases because the average distance between stimuli becomes smaller and the spatial relationships become more complex as more stimuli are added. To avoid this possible confound, the present work used simplified Chinese characters (see Fig. 1), always presented at fixation. These experiments used the same vocal responses as Hyman, (1953), whose method this work follows most closely.

Fig. 1 The eight simplified Chinese characters used in all three experiments

In Experiment 1, we establish that HHL is obeyed for our task. In Experiment 2, we separate the roles of stimulus and response uncertainty by manipulating the former while holding the latter constant. In Experiment 3, we verify that estimating the role of response uncertainty by subtracting the contribution of stimulus uncertainty from the simultaneous and confounded effect is appropriate. We also explore (and rule out) the possibility that all of our measures of these effects are contaminated by variations in the rate of exact repetitions of trials when the number of stimuli is manipulated.

There are four possible outcomes for this research. First, HHL could be entirely driven by stimulus uncertainty (as suggested by both Hick and Hyman). Second, HHL could be entirely driven by response uncertainty (as suggested by, e.g., Bricker). Third, both stimulus and response uncertainty could contribute equally to HHL, such that one bit of stimulus uncertainty is equivalent to one bit of response uncertainty. Finally, stimulus and response uncertainty could contribute differently, with one type of uncertainty playing a larger role, but both having some effect. This final possibility would imply that the effect of uncertainty is influenced by the ease of translating the stimulus into a response (response selection). As a preview of the results we find that stimulus and response uncertainty both contributes to HHL (although response uncertainty more so) and joins a variety of other effects that are best understood at the level of response selection (e.g., dual-task costs, selective attention effects, and stimulus–response compatibility effects). These findings also motivate a revision to the equation for Hick/Hyman Law (Eq. 1) as there are separate slope (*b*) parameters for stimulus- and response-based uncertainty.

General method

The three experiments all used the same general method. Following Hyman, (1953), participants first trained for six 1-h sessions using the full set of stimuli and responses for the given experiment (hereafter, this will be referred to as the base mapping for the experiment). During Sessions 7 and 9, the number of stimuli and/or responses was reduced (these are the reduced mappings). Session 8 was identical to Sessions 1–6 and was used to assess performance with the base mapping.

The stimuli were presented and the vocal responses were identified with PC computers using Microsoft's Visual Basic and voice-recognition software. The stimuli were easily and equally discriminable simplified Chinese characters (see Fig. 1) that measured 1.9 cm × 1.9 cm, presented in white on a black background. Participants sat 66 cm away from the screen, resulting in a visual angle of 1.65°. Because all of the stimuli appeared in the center of the screen, attention could be focused on a single location regardless of the number of possible stimuli. No participant reported any knowledge of written (or spoken) Chinese.

The responses were single syllable pseudo-words taken from Hyman, (1953): "bun," "boo," "bee," "bore," "by," "bix," "bev," and "bate" (i.e., the names of the first eight digits in English with the initial sound changed to /b/ to allow for equally accurate response-time measures). The base mapping of stimuli to responses was independently randomized for each participant.

Each trial began with the presentation of a fixation cross for 500 ms, followed by the stimulus at the center of the computer screen, which remained visible until the subject made a response or 3000 ms had elapsed. The inter-trial interval was 700 ms. Participants were instructed to respond as quickly and accurately as possible. Following incorrect responses, visual feedback was presented indicating both the participant's and the correct response.

Each block consisted of 32 trials with each possible stimulus presented an equal number of times in a pseudo-random order. Two randomly-selected warm-up trials were added to the beginning of every block (but omitted from the analysis). Each of the nine sessions included 20 blocks of which only the final 12 were included in the analysis. This was done to allow participants to adjust to the new tasks versions and ensure that they did not expect stimuli that were not included in the session. Trials that immediately followed an error (5–9 %, depending on experiment) were also omitted from the analysis (see, e.g., Rabbitt, 1966, for the rationale), as were response times below 100 ms or above 1250 ms (1 or 2 %, depending on experiment).

For the reported analyses, the data were further restricted to only those stimuli and responses that were common to all three conditions (see, Leonard, 1959, for the rationale). For example, in Experiment 1, one condition involved all eight stimuli and all eight responses, another involved only four of each, and the last only involved two of each. The critical analysis concerned only the data from trials using the two stimuli and two responses that were included in all three conditions. This was done in order to avoid comparing trials with different stimuli and different vocalizations; if different stimuli were more or less difficult to identify or if different responses were more or less difficult to produce, then including different stimuli and responses in different conditions would act as an unwanted confound.

Following the method of Hyman, (1953), separate regressions of mean RT onto condition were performed for each participant. Mean correlations were calculated by first converting the observed values of *r* to *z*-scores (using Fisher's transformation), finding the mean, and then converting back to a value of *r*. Mean regression coefficients were calculated in the usual manner, without transformation. Finally, standard errors were calculated using the method of Cousineau, (2005) as amended by Morey, (2008) such that they more accurately represent the error in estimating a within-subjects difference.

A total of 18 participants (6 per experiment, 11 females) were recruited from the Iowa City area, ages ranging from 19 to 26 years. Each participant was paid $10 per session and each participant completed at least four sessions per week (sometimes five) but never more than one session in a single day.

Experiment 1

Experiment 1 had two goals. The first was to replicate Experiment 1 of Hyman, (1953). On the likely assumption that HHL would be verified, the second goal was to acquire an estimate of total effect on mean RT for this task with simultaneous (i.e., confounded) changes in stimulus and response uncertainty. Experiment 1 included three conditions because the core of the phenomenon is the linear relationship between mean RT and average uncertainty (H), and we wished to minimize the number of conditions in order to maximize the amount of data per condition, while still being able to verify linearity.

Method

The base mapping consisted of all eight stimuli and all eight responses (hereafter: the *8:8 condition*), with each being equally likely, such that stimulus and response uncertainty were both 3 bits. One of the reduced mappings included only four of the stimulus/response pairs (the *4:4 condition*) with 2 bits each, and the other included only two of the pairs (the *2:2 condition*) with 1 bit each. Because each stimulus was mapped to a different response, stimulus uncertainty and response uncertainty were always the same (as they were in Hyman, 1953). The order of the two reduced mappings (run during Sessions 7 and 9) was counter-balanced across subjects.

Results and discussion

Response accuracy was very high, and there were no significant differences between the 2:2 (94.2 %), 4:4 (95.6 %), and 8:8 (95.4 %) conditions, $F < 1$, indicating that the differences in mean RT were not the result of a speed/accuracy tradeoff.

The mean RTs (with standard errors) for the critical test of HHL were 348 ± 16 ms, 451 ± 10 ms, and 497 ± 17 ms for the 2:2, 4:4, and 8:8 conditions, respectively (see Fig. 2 for plots of individual participants). The mean value of r was .95, ranging from .90 to .99 across subjects, which compares well with the mean r of .98 for Experiment 1 of Hyman, (1953). The mean slope of the function relating average uncertainty to mean RT (i.e., parameter b in Eq. 1) was 74 ± 12 ms/bit. This is much less steep than that observed by Hyman, (1953), but the difference is not unexpected. The slope of HHL is known to depend on stimulus–response compatibility (e.g., Dassonville et al., 1999; Smith, 1968a; Teichner & Krebs, 1974), and vocal responses to visual symbols enjoy much more set-level compatibility (Wang & Proctor, 1996) than vocal responses to spatial locations.

In summary, Experiment 1 verified that the present task produces data that obey HHL and provided an estimate of 74 ms/bit for the combined effects of stimulus and response uncertainty.

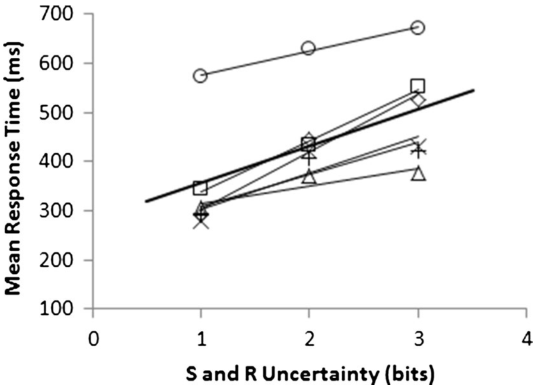

Fig. 2 Mean response time as a function of confounded stimulus and response uncertainty. (The mean trend is represented by the *bold line*, while *different symbols* indicate different individual participants.)

Experiment 2

The preferred way to separate the effects of two confounded manipulations is to repeat the experiment with one of the variables held constant. Experiment 2 does this by holding the number of responses constant at two, while including conditions with eight, four, or two stimuli. Thus, Experiment 2 provides a pure measure of the effects of stimulus uncertainty.

Method

The base mapping consisted of all eight stimuli, but only two responses (the *8:2 condition*), with a different pair of responses selected at random for each participant. Four stimuli were mapped onto each of the two responses and all stimuli occurred equally often. The base mapping was used in the six training sessions and session 8 as well. This produced a condition with 3 bits of stimulus uncertainty but only 1 bit of response uncertainty. One reduced mapping included only four of the stimuli, two for each of the two responses (the *4:2 condition*) and, therefore, had 2 bits of stimulus uncertainty and 1 bit of response uncertainty. The other reduced mapping used only one stimulus for each response (the *2:2 condition*) and had 1 bit of each type of uncertainty. As before, the order of the two reduced mappings was counter-balanced across subjects and occurred during Sessions 7 and 9.

Results and discussion

The mean accuracies were 96.5, 92.9, and 90.5 % in the 2:2, 4:2, and 8:2 conditions, respectively. There was a trend towards a difference between conditions, $F(2,10) = 3.46$, $\mathrm{MSE} = 0.002$, $\eta_p^2 = 0.409$, $p = .072$; however, this effect does not seem to indicate a speed-accuracy trade off as the lower accuracies were observed in the conditions with slower responses.

The mean RTs (with standard errors) for the 2:2, 4:2, and 8:2 conditions were 322 ± 20 ms, 333 ± 18 ms, and 372 ± 14 ms, respectively (Fig. 3), producing a mean slope of 25 ± 7 ms/bit and a mean r of .55. A restriction of range analysis (with respect to Experiment 1; see Hunter & Schmidt, 1990) results in a corrected r of .74, which is sufficient to use the results from the linear regression. The slope from Experiment 2 was significantly shallower than

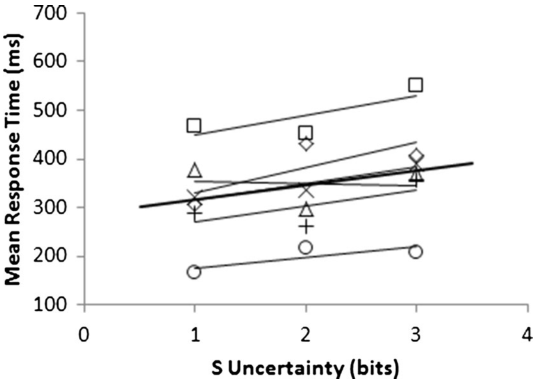

Fig. 3 Mean response times across different levels of stimulus uncertainty. The mean slope is represented by the *bold line* and individual participants are represented by *different symbols*

that from Experiment 1, $t(10) = 2.81$, Cohen's $d = 1.62$, $p = .018$, and was only marginally different from zero, $t(5) = 2.14$, Cohen's $d = 1.23$, $p = .086$.

It is possible that participants were thinking about the larger stimulus set even when using the reduced mappings, which would cause an underestimate the contribution of stimulus uncertainty (Fitts & Switzer, 1962). However, we note that this may occur in any experiment that reduces the number of stimuli and/or responses to estimate the slope, including those of Hick, (1952) and Hyman, (1953). We took steps to minimize this concern by explicitly instructing participants as to which S–R alternatives were possible during the block and testing the conditions on separate days, as well as providing many blocks of practice before collecting the data to be retained. Moreover, the mean RTs for the 2:2 conditions in Experiments 1 and 2 were similar (348 vs. 322 ms, respectively), suggesting that the participants in Experiment 2 were not expecting more stimuli than those in Experiment 1. Therefore, even if participants were unable to restrict their expectations, this does not appear to be driving the difference in slopes between the two experiments.

In summary, the combined results from Experiments 1 and 2 suggest that stimulus uncertainty accounts for approximately one-third (25 ms/bit out of 74 ms/bit) of the total slope in HHL. If the effects of stimulus and response uncertainty are assumed to be additive—which shall be tested in Experiment 3—then this also implies that response uncertainty is twice as important as stimulus uncertainty, at least for the present task.

Experiment 3

In the ideal situation, when attempting to separate the effects of two confounded manipulations, after conducting an experiment that holds one of the variable constants, the next experiment would hold the other variable constant. In the case of stimulus and response uncertainty, however, this is difficult to implement. In order to hold the number of stimuli constant while manipulating the number of responses, either the mapping of particular stimuli to responses would have to change at the same time as the number of responses (which would be a new confound) or different sets of stimuli would have to be used for each number-of-response condition (which would also be a new confound). And, yet, it does not seem appropriate to simply assume that the effects of stimulus and response uncertainty are additive, as is done when the slope for the effect of response uncertainty ($\approx$50 ms/bit) is estimated by subtracting the slope for stimulus uncertainty (25 ms/bit) from the slope for confounded stimulus and response uncertainty (74 ms/bit). Therefore, we devised an alternative approach.

The goal of Experiment 3 was to acquire separate measures of the effects of stimulus and response uncertainty without having to resort to a variable mapping or multiple stimulus sets. This was done by selecting three conditions, two of which differed only in terms of the number of stimuli (eight vs four, with four responses in both cases), and two of which differed only in terms of the number of responses (four vs two, with four stimuli in both cases). This allows for a direct comparison of the effects of the two types of uncertainty, as well as a replication of the specific slope values provided by the previous pair of experiments. The only weakness to this approach is the loss of any way to verify that the effects are both linear, because only two points are measured in each case. But, compared to the clear problems that would have been encountered using the other approaches (i.e., the introduction of new confounds that would then need to be separated), this seemed a minor loss.

Method

The base mapping included all eight stimuli, but only four of the responses (the *8:4 condition*), with a different set of responses selected for each participant. Two stimuli were mapped onto each of the responses and all stimuli occurred equally often. Again, this base mapping was used during the training sessions as well as session 8. This produced a condition with 3 bits of stimulus uncertainty and 2 bit of response uncertainty. One reduced mapping included only four of the stimuli, one for each of the four responses (the *4:4 condition*) entailing 2 bits of each type of uncertainty. The other reduced mapping also used only four of the eight stimuli, but this time, there were two stimuli for each of two responses (*the 4:2 condition*). This produced 2 bits of stimulus uncertainty and 1 bit of response uncertainty. As always, the order of the two reduced mappings was counter-balanced across subjects and occurred during Sessions 7 and 9.

The difference between the 8:4 and 4:4 conditions provides a pure estimate of the effect of stimulus uncertainty. Given the results from Experiment 2, this difference should be approximately 25 ms because the difference between these conditions is 1 bit of stimulus uncertainty (only) and the slope from Experiment 2 was 25 ms/bit. Conversely, the difference between the 4:4 and 4:2 conditions provides a pure estimate of effect of response uncertainty. Based on our previous findings and the simplifying assumption that stimulus uncertainty and response uncertainty will produce additive effects, this difference should be approximately 50 ms, since the difference here is 1 bit of response uncertainty and the difference between the slopes from Experiments 1 and 2 was 49 ms/bit. The subset of stimuli used for the two reduced conditions was randomly chosen at the outset of the experiment with the constraint that two of the stimuli would be shared across all the three conditions (i.e., 8:4, 4:4, and 4:2) because the analyses were restricted to the trials that are common across all conditions (see General Methods).

Results and discussion

The mean accuracies were 95.6, 97.5, and 95.8 % for the 4:2, 4:4, and 8:4 conditions, respectively. There was no significant difference between these values, $F < 1$.

The mean RTs for the 4:2, 4:4, and 8:4 conditions were 284 ± 9 ms, 345 ± 2 ms, and 369 ± 10 ms, respectively (see Table 1 for individual values). The estimate of the effect of stimulus uncertainty (obtained by comparing the 4:4 and 8:4 conditions) was 24 ± 10 ms/bit, which is highly consistent with the value from Experiment 2. The estimate of the effect of response uncertainty (based on the comparison of the 4:2 and 4:4 conditions) was 61 ± 2 ms/bit, which is slightly larger than the prediction produced via the difference between Experiments 1 and 2, but not significantly different from the predicted value, $t < 1$.

Because Experiment 3 provided measures of both effects from the same subjects, they can be directly compared with acceptable levels of power. The contribution of response uncertainty to HHL was significantly greater than that of stimulus uncertainty, $t(5) = 6.43$, Cohen's $d = 1.23$, $p = .001$. Taken together, these experiments indicate that response uncertainty is responsible for between two-thirds (49 out of 74) and three-quarters (61 out of 84) of the total effect of confounded manipulations of stimulus and response uncertainty.

Table 1 Mean RTs (in ms) for the three different mapping conditions for individual participants in Experiment 3

	Mapping condition		
Subject	8:4	4:4	4:2
1	372	349	281
2	566	505	409
3	296	307	254
4	383	357	302
5	271	254	205
6	326	298	244
Mean	369	345	284

Repetition effects

When average uncertainty is manipulated by changing the number of stimuli and/or number of responses, as was done here, the probability of a repetition (across adjacent trials) of the stimulus and/or response is also changed. As the number of stimuli, for example, is increased, the probability of a stimulus repetition decreases, likewise for responses. Given that stimulus/response repetitions are known to have profound effects on performance (e.g., Kornblum, 1969; Pashler & Baylis, 1991; Rabbit, 1968; Smith, 1968b), the possibility that some (if not all) of the supposed effects of stimulus and/or response uncertainty are actually due to the effects of these repetitions needs to be examined.

This issue was indirectly explored in Experiment 3 of Hyman, (1953), in which the sequential probabilities of stimuli and responses were manipulated. In that experiment, the same values of b were found when manipulating sequential probabilities of stimuli and responses as were found when varying the numbers of stimuli and responses and/or when varying the overall probabilities of the stimuli and responses. This finding suggests that uncontrolled repetitions do not have a major impact on the estimate of b. Kornblum, (1969) also examined the role that repetitions might play in the performance of tasks with varying numbers of stimuli and responses, noting that simpler tasks often involve more repetitions. However, neither of these studies (nor any others) included a direct test of the contribution of stimulus/response repetitions to the slope of HHL. Given that stimulus and response repetitions can have different effects on performance (e.g., Pashler & Baylis, 1991), taking a closer look at the issue would seem necessary before concluding that stimulus and response uncertainty have different effects on performance.

To do this, we compared the slopes and the values of r that were generated when stimulus/response repetitions are either included or excluded from the analysis. In practice, we removed all trials in which the response was the same as in the previous trial because this procedure eliminated trials in which the stimulus repeated and also eliminated trials in which the stimulus changed but the response repeated. In Experiment 1, including all of the data produced a mean slope (with standard error) of 74 ± 12 ms/bit with a mean *r* of .95, while excluding repetitions produced a mean slope of 111 ± 16 ms/bit with a mean *r* of .98. In Experiment 2, including all of the data produced a slope of 25 ± 7 ms/bit with a mean *r* of .55 (uncorrected), while excluding repetitions produced 43 ± 12 ms/bit with a mean *r* of .60. As can be seen, the removal of repetitions serves to increase the magnitude of the uncertainty effect. Even more, given that the effect of excluding repetitions was numerically larger in Experiment 1 than in Experiment 2, the relative roles of stimulus and response uncertainty were possibly even more lopsided in favor of the latter than the reported analyses would suggest.

We also evaluated the role of repetition benefits in Experiment 3. The repetition benefits for the 8:4 mapping was 63 ± 5 ms, for the 4:4 mapping was 53 ± 8 ms, and for the 4:2 mapping was 36 ± 10 ms. As with overall mean RT, the repetition benefits were larger when the change was 1 bit of response uncertainty (4:4 compared to 4:2) than when the change was 1 bit of stimulus uncertainty (8:4 compared to 4:4). Finally, mean RTs for the three conditions with repetitions removed were 375 ± 7 ms, 363 ± 4 ms, and 301 ± 11 ms for the 8:4, 4:4, and 4:2 mappings, respectively, again finding a much larger effect for response than stimulus uncertainty. In sum, response uncertainty has consistently larger effects on mean RT than stimulus uncertainty, regardless of whether repetitions are retained or removed.

General discussion

Hick/Hyman Law (Hick, 1952; Hyman, 1953) holds that mean response time is a linear function of average uncertainty (see Eq. 1, above). The goal of the present work was to determine the separate roles that stimulus and response uncertainty play in determining the slope (*b*) of HHL. We found that in a visual-vocal task after moderate practice, between two-thirds and three-quarters of the slope was accounted for by response uncertainty. We also verified that inter-trial repetition effects do not account for HHL as there was still a high degree of linearity when all repetitions were removed from the analysis.

These findings motivate an expansion of the existing equation that is used to express HHL. The amended version of the law would include separate measures of stimulus and response uncertainty, and separate slope parameters for their effects:

$$\text{mean RT} = a + b_s \mathrm{H}_s + b_r \mathrm{H}_r, \qquad (5)$$

where H_s is average stimulus uncertainty, H_r is average response uncertainty, and a, b_s, and b_r are constants estimated by regression, with b_s representing the effect of stimulus uncertainty (in ms/bit) and b_r representing the effect of response uncertainty (also in ms/bit). The main justification for dividing what used to be one measure of uncertainty and one estimated slope into two separate measures of uncertainty and two separate slopes is the finding that a bit of stimulus uncertainty contributes much less to performance than a bit of response uncertainty. To accommodate this, separate values and slopes are needed within the equation.

Because Experiment 3 estimated the contribution of stimulus uncertainty with conditions that used a different number of responses (8:4 vs. 4:4) than the conditions used in Experiment 2 (8:2 vs. 4:2 vs. 2:2), the similarity across the three experiments also suggests that the formulation encapsulated by Eq. 5 does not require an interaction term. Thus, at least within the narrow range of values examined here, which is the range for which HHL is most often and widely applied, the effects of stimulus uncertainty do not depend on the number of responses, and the effects of response uncertainty do not depend on the number of stimuli. In sum, changes in the stimuli, the responses, and the amount of practice may each affect the relative contributions of stimulus uncertainty and response uncertainty to the observed slope, but the present findings demonstrate that the two forms of uncertainty both contribute to the slope and that their contributions are not necessarily equal.

Implications for models of Hick/Hyman Law

Two broad classes of models have been proposed as explanations of HHL: stimulus-driven accounts and response-driven accounts. Stimulus-driven accounts hold that the effects of uncertainty arise during stimulus processing. Both Hick, (1952) and Hyman, (1953) focused on the observer's uncertainty with regard to the stimulus (e.g., "reaction time as a function of the number of stimulus alternatives"; Hyman, 1953, p. 191), and hardly mention the responses, even though the number of responses was always the same as the number of stimuli. Hick, (1952) suggested that the logarithmic relationship was produced by a series of hierarchical decisions, each dividing the number of remaining possibilities into half, and that these decisions concerned the identity of the stimulus (as opposed to the required response). In contrast, an account more consistent with the present results was provided by Lacouture, Li, & Marley, (1998), who modeled reaction times for categorizing unidimensional stimuli into varying numbers of response categories. Although this task requirement gives rise to various phenomena not observable in typical Hick-Hyman tasks (e.g., the bow effect and the category boundary effect), the number of response categories played a much larger role in determining RT than the number of possible stimuli.

More recently, Schneider & Anderson, (2011) proposed a memory-based model of response selection in which the HHL arises from a weakening of activation from contextual cues that prime the possible stimulus–response (S–R) associations. In essence, the HHL stems from the "fan effect" according to this model (Anderson & Reder, 1999): the more S–R associations included in a task-set, the less each association is activated before the presentation and processing of the stimulus; the less pre-activation prior to the onset of the stimulus, the longer is the RT. If it is assumed that each stimulus activates a distinct S–R association (or that each S–R association has a unique triggering stimulus), then the number of possible stimuli should play the critical role in determining mean RT. The present finding that response uncertainty plays a greater role than stimulus uncertainty in determining mean RT is inconsistent with this type of model.

Alternatively, response-based accounts stress the competition between different response options (e.g., Berlyne, 1957a, b; Usher & McClelland, 2001; Usher et al., 2002) and, therefore, can readily explain the larger role of response uncertainty. For example, Usher & McClelland, (2001); see also, Usher et al., 2002; Hawkins et al., 2012; Lacouture et al., 1998) modeled response selection as a set of accumulators gathering evidence for each of the possible responses. The larger the number of possible responses, the more likely random noise will produce supra-threshold activation in at least one of the incorrect alternatives. One way to prevent an increase in error rates with more response alternatives is to raise the threshold so that more activation is required to trigger the response (Usher & McClelland, 2001). Such models are highly consistent with the present finding that response uncertainty can play the larger role in determining mean RT, although they will also need to be amended to account for the separate effect of stimulus uncertainty.

Implications for models of response selection

Converging evidence for a response accumulator mechanism can be found in a study examining similarity-based learning effects (Wifall, McMurray, & Hazeltine, 2014). In these experiments, participants learned to produce either similar or dissimilar three-finger key-presses (analogous to how a pianist plays a chord) and received nearly identical amounts of practice to the participants in the present study. Critically, chords cued by dissimilar stimuli were produced faster after training than chords cued by similar stimuli (even though no difference was observable early in training) regardless of the stimuli used during training. The authors concluded that the difference between the similar and dissimilar chords emerged with practice because the cues for similar chords activated multiple learned responses whereas the cues for the dissimilar chords activated only the single learned response. Thus, competition among response options increased with training. Because the similarity effect depended on the type of cue used during the test but not on the type of cue used during training, it appeared that the learning involved the encoding of response representations that could be activated by a range of stimuli. This conclusion is consistent with the present findings in that it suggests competition among response alternatives plays a key role in RT after moderate practice, as in the present study. If we assume that a stimulus activates response competitors to some degree in addition to the correct response, then the more competitors, the greater is the competition. According to this account, when the S-R mappings are highly compatible, a stimulus will activate competitors to a lesser degree, and the slope of function relating the number of alternatives and RT will decrease.

Competition among response alternatives provides a straightforward explanation for the effects of response uncertainty, but it is not sufficient to accommodate the small effect of stimulus uncertainty observed in the present experiments. One possible solution is to posit competition in both perceptual and response processing. Alternatively, models in which response features prime perceptual features (e.g., Metzker & Dreisbach, 2009; Müsseler & Hommel, 1997) may start to explain the present pattern of results. What is clear is that the number of possible responses can have a strong effect on RT, arguing against models in which stimuli trigger responses in a strictly feed-forward fashion.

It is also possible that the number of stimuli affects the degree to which a given stimulus activates response competitors. Each practice trial may reduce the extent to which the presented stimulus activates the competitors and, therefore (provided the stimuli are dissimilar, see Wifall et al., 2014), the smaller the number of stimuli, the greater is the reduction in the activation of competitors for a given number of practice trials.

More generally, these findings join a larger literature that demonstrates performance is greatly affected by the ease of translating the stimulus into the response. In other words, the key "job" for the decision system is to decide which response to make; the identity of the stimulus is only relevant because it (usually) provides evidence that is germane to this decision. This contrasts with both Hick's and Hyman's focus on how stimulus information affects performance, as illustrated by the title of Hyman's, (1953) paper, "Stimulus information as a determinant of reaction time" (emphasis added). This is not surprising as most of the research from early information processing was directed at the identification processes and not response selection.

More recently, a variety of effects have been best understood as the result of processing at the level of response selection. For example, dual-task costs (e.g., Pashler, 1984), selective attention effects (e.g., Eriksen & Eriksen, 1974; Simon, 1969; Stroop, 1935) and stimulus–response compatibility effects (e.g., Fitts & Seeger, 1953) have all been theorized to result from changes in the duration of the response selection stage. That is, these phenomena are all generally thought to reflect the increased difficulty of resolving conflict among competing response options rather than increased difficulty in identifying the stimulus. Such a conclusion is clearly consistent with the present account of HHL that emphasizes response uncertainty. Thus, these new findings contribute to body of research indicating that response selection for even simple tasks involves more than stimuli activating the appropriate responses in a strictly feed-forward manner. Additional work with other sets of stimuli and responses and different levels of practice may reveal situations where stimulus uncertainty plays a larger role in the determination of RT than response uncertainty. However, the present results make clear that the two forms of uncertainty can play different roles, and separate components are needed for the HHL equation.

Acknowledgments The authors wish to thank Robert Proctor and three anonymous reviewers for their helpful comments on earlier versions of this manuscript. This project was supported by NIH Grant R03 DA031583-01A1 awarded to EH.

References

Anderson, J. R., & Reder, L. M. (1999). The fan effect: new results and new theories. *Journal of Experimental Psychology: General, 128*, 186–197.

Attneave, F. (1959). *Applications of information theory to psychology: a summary of basic concepts, methods, and results*. New York: Holt-Dryden.

Berlyne, D. E. (1957a). Conflict and choice time. *British Journal of Psychology, 48*, 106–118.

Berlyne, D. E. (1957b). Uncertainty and conflict: a point of contract between information-theory and behavior-theory concepts. *Psychological Review, 64*, 329–339.

Bricker, P. D. (1955). The identification of redundant stimulus patterns. *Journal of Experimental Psychology, 49*, 73–81.

Cousineau, D. (2005). Confidence intervals in within-subject designs: a simpler solution to Loftus and Massons's method. *Tutorials in Quantitative Methods for Psychology, 1*, 42–45.

Dassonville, P., Lewis, S. M., Foster, H. E., & Ashe, J. (1999). Choice and stimulus-response compatibility affect duration of response selection. *Cognitive Brain Research, 7*, 235–240.

Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception and Psychophysics, 16*, 143–149.

Fitts, P. M., & Seeger, C. M. (1953). S-R compatibility: spatial characteristics of stimulus and response codes. *Journal of Experimental Psychology, 46*, 199–210.

Fitts, P. M., & Switzer, G. (1962). Cognitive aspects of information processing: I. The familiarity of S-R sets and subsets. *Journal of Experimental Psychology, 63*, 321–329.

Hawkins, G., Brown, S. D., Steyvers, M., & Wagenmakers, E.-J. (2012). Context effects in multi-alternative decision making: empirical data and a Bayesian model. *Cognitive Science, 36*, 498–516.

Hick, W. E. (1952). On the rate of gain of information. *Quarterly Journal of Experimental Psychology, 4*, 11–26.

Hunter, J. E., & Schmidt, F. L. (1990). *Methods of meta-analysis. Correcting error and bias in research findings*. London: Sage Publications.

Hyman, R. (1953). Stimulus information as a determinant of reaction time. *Journal of Experimental Psychology, 30*, 188–196.

Jamieson, R. K., & Mewhort, D. J. K. (2009). Applying an exemplar model to the serial reaction-time task: anticipating from experience. *The Quarterly Journal of Experimental Psychology, 62*, 1757–1783.

Keele, S. W. (1970). Effects of input and output modes on decision time. *Journal of Experimental Psychology, 85*, 157–164.

Kornblum, S. (1969). Sequential determinants of information processing in serial and discrete choice reaction time. *Psychological Review, 76*, 113–131.

Kveraga, K., Boucher, L., & Hughes, H. (2002). Saccades operate in violation of Hick's law. *Experimental Brain Research, 146*, 307–314.

LaBerge, D., Legrand, R., & Hobbie, R. K. (1969). Functional identification of perceptual and response biases in choice reaction time. *Journal of Experimental Psychology, 79*, 295–299.

LaBerge, D., & Tweedy, J. R. (1964). Presentation probability and choice time. *Journal of Experimental Psychology, 68*, 477–481.

Lacouture, Y., Li, S.-C., & Marley, A. A. J. (1998). The roles of stimulus and response set size in the identification of categorization of unidimensional stimuli. *Australian Journal of Psychology, 50*, 165–174.

Leonard, J. A. (1959). Tactual choice reactions. *Quarterly Journal of Experimental Psychology, 11*, 76–83.

Longstreth, L. E. (1988). Hick's law: its limit is 3 bits. *Bulletin of the Psychonomic Society, 26*, 8–10.

Longstreth, L. E., El-Zahhar, N., & Alcorn, M. B. (1985). Exceptions to Hick's law: explorations with a response duration measure. *Journal of Experimental Psychology: General, 114*, 417–434.

Metzker, M., & Dreisbach, G. (2009). Bidirectional priming processes in the Simon task. *Journal of Experimental Psychology: Human Perception and Performance, 35*, 1770–1783.

Morey, R. D. (2008). Confidence intervals from normalized data: a correction to Cousineau (2005). *Tutorial in Quantitative Methods for Psychology, 4*, 61–64.

Müsseler, J., & Hommel, B. (1997). Blindness to response-compatible stimuli. *Journal of Experimental Psychology: Human Perception and Performance, 23*, 861–872.

Pashler, H. (1984). Processing stages in overlapping tasks: evidence for a central bottleneck. *Journal of Experimental Psychology: Human Perception and Performance, 10*, 358–377.

Pashler, H., & Baylis, G. (1991). Procedural learning: 2. Intertrial repetition effects in speeded-choice tasks. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 17*, 33–48.

Pollack, I. (1959). Message uncertainty and message reception. *The Journal of the Acoustical Society of America, 31*, 1500–1508.

Rabbit, P. M. A. (1968). Repetition effects and signal classification strategies is serial choice-response tasks. *The Quarterly Journal of Experimental Psychology, 20*, 232–240.

Rabbitt, P. M. A. (1959). Effects of independent variations in stimulus and response probability. *Nature, 183*, 1212.

Rabbitt, P. M. A. (1966). Errors and error corrections in choice-response tasks. *Journal of Experimental Psychology, 71*, 264–272.

Schneider, D. W., & Anderson, J. (2011). A memory-based model of Hick's law. *Cognitive Psychology, 62*, 193–222.

Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Technical Journal, 27*(379–423), 623–656.

Shannon, C. E., & Weaver, W. (1949). *The mathematical theory of communication*. Urbana: Univ. of Illinois Press.

Simon, J. R. (1969). Reactions towards the source of stimulation. *Journal of Experimental Psychology, 81*, 174–176.

Smith, E. E. (1968a). Choice reaction time: an analysis of the major theoretical positions. *Psychological Bulletin, 69*, 77–110.

Smith, M. C. (1968b). Repetition effect and short-term memory. *Journal of Experimental Psychology, 77*, 435–439.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology, 18*, 643–662.

Teichner, W. H., & Krebs, M. J. (1974). Laws of visual choice reaction time. *Psychological Review, 81*, 75–98.

Usher, M., & McClelland, J. L. (2001). The time course of perceptual choice: the leaky, competing accumulator model. *Psychological Review, 108*, 550–592.

Usher, M., Olami, Z., & McClelland, J. L. (2002). Hick's law in a stochastic race model with speed-accuracy tradeoff. *Journal of Mathematical Psychology, 46*, 704–715.

Wang, H., & Proctor, R. W. (1996). Stimulus-response compatibility as a function of stimulus code and response modality. *Journal of Experimental Psychology: Human Perception and Performance, 22*, 1201–1217.

Welford, A. (1968). *The fundamentals of skill*. London: Methuen.

Wifall, T., McMurray, B., & Hazeltine, E. (2014). Perceptual similarity affects the learning curve (but not necessarily learning). *Journal of Experimental Psychology: General, 143*, 312–331.

Wright, C. E., Marino, V. M., Belovsky, S. A., & Chubb, C. (2007). Visually-guided, aimed movements can be unaffected by stimulus—response uncertainty. *Experimental Brain Research, 179*, 475–496.