E C O N O M I C S   B U L L E T I N

Guessing and gambling

Albert Burgos
Universidad de Murcia

Abstract

Scoring methods in multiple−choice tests are usually designed as fair bets, and thus random guesswork yields zero expected return. This causes the undesired result of forcing risk averse test−takers to pay a premium in the sense of letting unmarked answers for which they have partial but not full knowledge. In this note I use a calibrated model of prospect theory [Tversky and Kahneman (1992, 1995))] to compute a fair rule which is also strategically neutral, (i.e. under partial knowledge answering is beneficial for the representative calibrated agent, while under total uncertainty it is not). This rule is remarkably close to an old rule presented in 1969 by Traub et al. in which there is no penalty for wrong answers but omitted answers are rewarded by 1/M if M is the number of possible answers.

I thank Maria F. Morales, who has provided useful feedback on this paper.
Citation: Burgos, Albert, (2004) "Guessing and gambling." *Economics Bulletin,* Vol. 4, No. 4 pp. 1−10
Submitted: January 30, 2004. **Accepted:** February 28, 2004.
URL: http://www.economicsbulletin.com/2004/volume4/EB−04D80001A.pdf

"Taking the GRE is a game with its own rules, traps, and measures of success...How you do on the GRE is an indication of how well you play the game..." (Jacobson, 1993 p. 27)

1 Introduction

Multiple-choice tests are often used as an easy and objective way to rank people in examinations. Examples go from academic exams to achievement or aptitude tests as the SAT (Scholastic Aptitude Test), or the GRE (Graduate Record Examinations). Beyond the academic world, a quick search on the internet will show a plethora of organizations and software solutions that can provide on-line financial consulting services which include recruitment tests.

Because the good or bad design of multiple choice tests can have such life-altering consequences for students, companies, and jobseekers, countless scholars (mostly psychologists or educators) have dealt extensively with the development and comparison of different scoring rules. In particular the random guessing problem, i.e. the existence of a positive probability of selecting correct answers to items about which the test taker knows nothing, has inspired a voluminous theoretical and empirical research.[1] About 80 years ago scholars introduced a correction-for-guessing feature, which consists in the use of scoring formulas under which the respondent's expected score for an item is the same whether he omits it or picks one answer from the set of options following a uniform probability distribution. Implicit in this approach it was assumed that (i) test takers were expected score maximizers, and (ii) with respect to each item, an examinee either was in a state of absolute certainty (i.e. a 100% sure of knowing the correct answer) or in a state of total uncertainty (hence, assigning equal subjective probability to each answer). Insofar as these two assumptions has been repeatedly challenged on empirical grounds, the so-called correction for guessing has not solved the problem that motivated it.[2] If omitting is a dominated strategy when the scoring rule has no correction for guessing, facing a guessing correction may discourage answering in situations in which the respondent has partial knowledge but he is risk averse. As Budescu and Bar-Hillel (1993) put it:

> *"Drawing the line between the kind of guessing that should be encouraged (e.g. mining partial knowledge) and the kind that should, perhaps, not be (e.g. capitalizing on chance) is very difficult. Moreover, test takers who choose to disregard this distinction cannot be prevented from doing so" (p. 288).*

This paper was the first one to suggest only a decade ago a decision-theory analysis of the problem of guessing in tests. Recently, Bereby-Meyer *et al.* (2002) have offered the first experimental

[1] This pure random guessing inflates the measurament error of tests, reducing their reliability.

[2] This theoretical debate has had practical consequences: SAT exam changed its scoring rule from no penalization to penalization of wrong answers in 1953, whereas GRE exam took the opposite way in 1984.

evidence on the existence of framing effects in test taking, a result they link to examinees' choices following prospect theory (Tversky and Kahneman (1992), Kahneman and Tversky (1979)). In this letter I go one step further on their appeal to framing effects by representing test takers as using a parametrized version of prospect theory, and choosing parameters so that test takers' preferences match real behavior in risky situations both in the laboratory and in real economic applications.

Guessing analysis using these 'calibrated' subjects singles out a rule which is 'fair' both from a statistical viewpoint (by offering a zero expected score for random guesswork) and is also neutral from a strategic viewpoint, by leaving calibrated agents indifferent between guessing at random or omitting.[3] Quite remarkably, for tests having 4 or 5 options per item this neutral rule almost coincides with an old rule proposed in Traub *et al.* (1969) which, unfortunately, no major testing program employs it. Of course, as more theoretical work and empirical evidence of this kind accumulates, the situation may change.

2 To guess or not to guess

When facing an item in a multiple-choice test an examinee must choose wether to pick an answer among M possible options (and perhaps risk losing points if the chosen answer turns out to be incorrect) or to omit it, (a riskless option). In what follows, I shall restrict my attention to items with a unique correct answer. Formally, let C_i $(i = 1, \ldots, M)$ denote the event "the correct answer is placed in the i^{th} position". Thus, $C_1, \ldots, C_M$ are disjoint events partitioning the set of states of nature S. Events $C_1, \ldots, C_M$ are mapped into judged probabilities $p_1, \ldots, p_M$ where p_i reflects the probability that the examinee assigns to the event C_i. Probability judgements need not be additive, but they do satisfy *subadditivity*, which implies $p_1 + \cdots + p_M \geq 1$, and *binary complementarity*, which implies that the event NC_i (denoting the complementary of C_i on S) is judged to have probability $1 - p_i$ (see Fox and Tversky (1998) and Wu and González (1999), among others). If x is the reward associated with a right response, y with an omission and z with a wrong response $(x > y > z)$, the examinee is therefore facing the following decision tree:

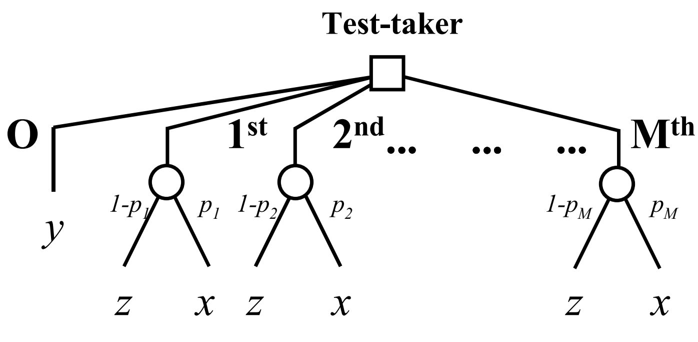

FIGURE 1A: One-person game in test taking.

[3] And therefore drawing the line between capitalizing on chance and having partial knowledge, as Budescu and Bar-Hillel asked for.

Without loss of generality, I shall assume that there is one option which is judged to be the most likely true. Any test taker whose preferences are monotone with respect to the relation of first order stochastic dominance will find strategically equivalent the above tree to the reduced tree below. For simplicity, subindexes indicating the option chosen are suppressed.

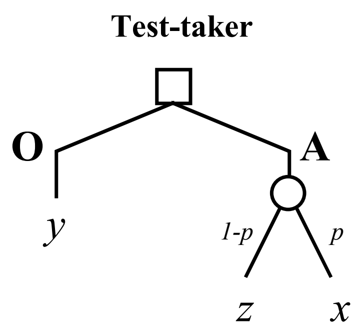

FIGURE 1B: Reduced one-person game in test taking.

The examinee is required to choose either (**A**) which represents picking the answer with the highest probability of being correct or (**O**) meaning to omit the question. Branch **A** yields therefore the prospect $(x, C; z, NC)$ offering a subjective chance p of event C, or $1-p$ chance of NC, and branch **O** yields y with certainty. In terms of outcomes and judged probabilities, if we set the outcome of a correct answer, x, to be equal to 1, different scoring rules are characterized by y $(0 \leq y < 1)$, the payoff associated with the omission of items, and z $(z < y)$, the penalization for wrong answers. For instance, the GRE general exam in the United States uses the *Number of Right* (here denoted NR) rule, which sets $y = z = 0$, whereas the SAT examinations have been employing the *Standard Correction* (here SC) rule, for which $y = 0$ and $z = -1/(M-1)$ where $M \geq 2$ is the number of possible answers for each item. A rational test maker endowed with risk preferences represented by V will choose **A** whenever the value under V of the risky prospect $\mathcal{A} = (x, p; z, 1-p)$ exceeds the value of the degenerate prospect $\mathcal{O} = (y, 1)$, even if knowledge is not perfect and answering involves some partial guessing.[4] The profitability of guessing is therefore expressed as

$$V(\mathcal{A}) - V(\mathcal{O}) = V(x, p; z, 1-p) - V(y)$$

whose value depends on the utility theory employed.

The formulation of risk preferences in this article adopts the theoretical framework of *Cumulative Prospect Theory* (CPT for short) in which gains and losses, rather than final outcomes are considered the carriers of value. CPT is a fairly general theory that lends itself very readily to an examination of guessing in multiple-choice tests. It assumes a continuous strictly increasing value function defined over outcomes satisfying $v(0) = 0$, and two weighting functions,

[4] A fact that is not always understood in education research. When penalization for wrong answers was first introduced in the SAT, examinees were simply instructed not to guess at all.

w^+ and w^- for events leading to gains and losses respectively which are nondecreasing and satisfy $w^+(0) = w^-(0) = 0$ and $w^+(1) = w^-(1) = 1$. According to CPT, whenever $z \leq 0$, prospect $\mathbf{A}$ can be written as the mixture of its positive part, $\mathbf{A}^+ = (1, p; 0, 1-p)$ for which $V(\mathcal{A}^+) = w^+(p)\,v(1) + [w^+(1) - w^+(p)]\,v(0)$, and its negative part, $\mathcal{A}^- = (0, p; z, 1-p)$, for which $V(A^-) = w^-(1-p)\,v(z) + [w^-(1) - w^-(1-p)]\,v(0)$. The value of prospect $\mathcal{A}$ is then $V(\mathcal{A}) = V(A^+) + V(\mathcal{A}^-)$ and therefore the condition for guessing is:

$$V(\mathcal{A}) - V(\mathcal{O}) > 0 \Longleftrightarrow w^+(p)\,v(1) + w^-(1-p)\,v(z) - v(y) > 0 \tag{1}$$

In order to accommodate a pattern of choices exhibiting risk seeking for gains and risk aversion for losses of low probability combined with risk aversion for gains and risk seeking for losses of high probability (the so-called fourfold pattern of risk attitudes[5]), Tversky and Kahneman suggested (i) a value function that is concave for gains, convex for losses and steeper for losses than for gains, and (ii) S-shaped weighting functions both for gains and for losses, so that small probabilities are overweighted, whereas moderate and high probabilities are underweighted.

In accordance with the homogeneity of preferences observed in experimental research, Tversky and Kahneman derive a two-part power function of the form

$$v(x) = \begin{cases} x^\alpha & \text{if } x \geq 0, \\ -\lambda(-x)^\beta & \text{if } x < 0, \end{cases}$$

where λ controls the effect of loss aversion, and can assume any positive value. The greater the value of λ, the more pronounced loss aversion. In addition, parameters α and β control respectively for the sensitivity of value to the size of positive and negative outcomes.

Tversky and Kahneman also suggest weighting functions for gains and losses that are first concave and then convex in p (overweighting low, and underweighting high probabilities). They calibrate their model against experimental data using the following functional form, also suggested in Camerer and Ho (1991, 1994):

$$w^+(p) = \frac{p^\gamma}{[p^\gamma + (1-p)^\gamma]^{1/\gamma}} \quad \text{and} \quad w^-(p) = \frac{p^\delta}{\left[p^\delta + (1-p)^\delta\right]^{1/\delta}}.$$

The calibration of the general five-parameter model for individual subjects appears in Tversky and Kahneman (1992). The median values taken from their work appear in Table 1 below. Under the additional assumption $\gamma = \delta$, Camerer and Ho (1991) give their own mean estimate of this weighting function using different laboratory data. Bradley (2003) also assumes $\gamma = \delta$ and offers mean estimates of the ratios γ/α and γ/β using data from betting behavior in horse races.[6] Table 1 summarizes these estimates.

---

[5] See, for instance, Tversky and Wakker (1995) and the references therein.

[6] Jullien and Salanié (2000) provide a different calibration of CPT using data on gamblers at horse racing tracks. Their model doesn't allow for an estimation of the value function for losses, and therefore concentrate on how

TABLE 1—CALIBRATIONS OF THE CPT MODEL

Parameters:	λ	α	β	γ	δ	γ/α	γ/β	g	h
Tversky and Kahneman:	2.25	0.88	0.88	0.61	0.69	-	-	-	-
Camerer and Ho ($\gamma = \delta$):	-	-	-	0.56	-	-	-	-	-
Bradley ($\gamma = \delta$):	-	-	-	-	-	0.818	0.68	-	-

Considering that the data of these studies come from different sources and methodologies, the values obtained are remarkably close. Notice that the estimates of Bradley and Camerer and Ho together suggest a value of α around 0.69, and for β around 0.82. Thus the estimates of Tversky and Kahneman seem quite robust to the use of field data instead of laboratory choices. This paper avoids the estimation problem altogether by utilizing the parameter set of Tversky and Kahneman. Given these parameter values, it is possible to compute the value of guessing $V(\mathcal{A}) - V(\mathcal{O})$ for any scoring rule

3 To correct or not to correct for guessing? And if so, how?

If the question "To guess or not to guess" must be faced by test takers, test makers must face the question "To correct or not to correct for guessing?" In the no-correction option, the test maker offers examinees a dominant strategy and thus removes the penalty against more risk averse test takers. If test makers expect examinees to be heterogeneous in their risk attitudes (because of cultural background, for instance) it might be reasonable to settle for the NR rule and choosing $y = z = 0$. If, however, examinees can be pictured as a single representative test taker holding calibrated preferences, test makers should choose y and z in order to minimize the incentives for pure random guesswork without hindering guessing in situations of partial knowledge.

Notice from inequality (1) that, in the context of CPT, the profitability of guessing depends positively upon $w^+(p)$ and negatively on the degree of loss aversion. Thus, in the NR rule ($y = z = 0$), the value of guessing coincides with $w^+(p)$. Since in the calibrated model $w^+(1/M) > 1/M$ for all integer $M > 2$, the value of pure random guessing under CPT is higher than under expected utility theory (assuming equal value function over outcomes).[7]

In the SC rule ($y = 0$, $z = -1/(M-1)$), the value of $V(\mathcal{A}) - V(\mathcal{O})$ is always negative for

probabilities of gains are weighted differently from probabilities of losses. They find no evidence that weighting functions change from concave to convex and use simple power functions $w^+(p) = p^g$ and $w^+(p) = p^h$ to find a slightly convex weighting function for gains and a concave function for losses. When used with Bradley's data, Jullien and Salanié's estimations ($g = 1.162$ and $h = 0.318$), yield a convex value function both for gains and for losses, in sharp contrast with the other studies. Therefore, their results seem to be too context-specific for the purpose of this paper.

[7] This result holds in general for risk preferences satisfying *lower subadditivity*, which implies a concave weighting function for low probabilities. There is widespread experimental evidence supporting this property, which implies optimism for unlikely events (see Wakker (2001)). The reverse inequality holds, however, if $M = 2$.

all relevant values of M.[8] (see the plot in figure 2 below). Thus, the overweighting of small probabilities is more than compensated by the strong loss aversion suggested by Tversky and Kahneman's calibration. In fact, loss aversion is so strong that omitting is the optimal strategy even for quite high judged probabilities of success. For instance, for a typical value of the number of options, $M = 4$, the lowest value of p that makes guessing profitable is $p = 0.469$, which is almost twice the objective chance of being right. The risk premium paid by calibrated agents is then close to 90% of the objective chance of success. This seems too much risk premium for a reasonable test. For other typical values of M the minimum judged probability that makes guessing profitable always exceeds in more that 80% the value of the objective success probability (see table 2 below).

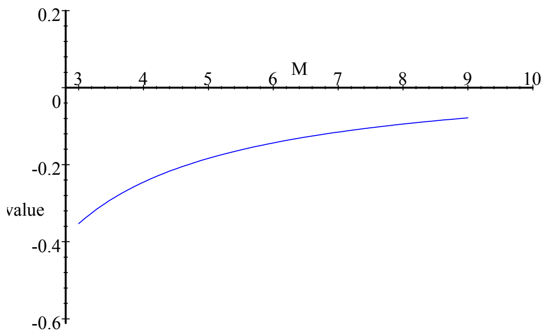

FIGURE 2: The value of pure guessing under the SC rule.

.

TABLE 2—MINIMUM JUDGED PROBABILITY FOR CHOOSING ANSWERING OVER OMITTING (CALIBRATED SUBJECTS)

M (Objective probability $1/M$)	2 (0.500)	3 (0.333)	4 (0.250)	5 (0.200)	6 (0.166)
Standard Correction rule:	0.788	0.603	0.469	0.372	0.302

Table 2 says that adopting the SC rule as a way to avoid random guessing leads to unacceptable costs imposed on the representative test taker, hindering his ability to make partial guesses. This brings us back to the problem posed by Budescu and Bar-Hillel. Is there a *fair* rule incentivating guessing if and only if the examinee has partial knowledge? By 'fair', I mean a scoring rule for which one's expected score is the same wether one guesses at random or omits, and thus returns a expected score of zero to someone who knows nothing at all. The necessary condition for fairness is therefore $1/M + z(M-1)/M - y = 0$, yielding

$$z = \frac{M}{M-1}\left(y - \frac{1}{M}\right).$$

[8] In fact, for the calibrated parameters, it becomes positive only when $M > 27$. Thus, no SC test in real life offers incentives for pure random guessing.

A scoring rule will be *strategically neutral* if guessing in a situation of partial knowledge is a dominant strategy for the representative test taker whereas random guessing it is not. According to this terminology, Budescu and Bar-Hillel were asking for a scoring rule both fair and strategically neutral. Neither the NR rule nor the SC rule are strategically neutral. The NR rule offers guessing as a dominant strategy for any test taker with monotonic preferences, whereas in the SC rule the representative test taker with calibrated preferences would choose omitting in a question for which his judged probability of success lies between $1/M$ and the minimum value that makes guessing profitable (see again table 2).

Of course, within the expected utility framework there are no fair rules which are also strategically neutral unless we allow the representative test taker to have a utility function with both concave and convex regions. Within the CPT framework, however, the nonlinear transformation of the probability scale leaves room for choosing a reward y and a penalty z implementing a rule which is are both fair and strategically neutral Notice that for CPT-examinees a strategically neutral rule must satisfy

$$w^{+}\left(\tfrac{1}{M}\right)v\left(1\right)+w^{-}\left(\tfrac{M-1}{M}\right)v\left(\tfrac{M}{M-1}\left(y-\tfrac{1}{M}\right)\right)-v(y)=0 \quad \text{if } y<1/M$$

$$w^{+}\left(\tfrac{1}{M}\right)+\left(1-w^{+}\left(\tfrac{1}{M}\right)\right)v\left(\tfrac{M}{M-1}\left(y-\tfrac{1}{M}\right)\right)-v(y)=0 \quad \text{if } y\geq 1/M.$$

Plugging Tversky and Kahneman's calibrated parameters in, we get for any given value of $M>2$ the exact value of y which solves the condition above.[9]

Table 3—Reward for omitting in a fair and strategically unbiased rule (calibrated subjects)

M (Objective probability $1/M$)	2 (0.500)	3 (0.333)	4 (0.250)	5 (0.200)	6 (0.166)
fair and strategically neutral rule :	-	0.882	0.257	0.191	0.149

Since most tests offer at least four choices per item, the rewards for omitting that make a fair rule strategically neutral are remarkably close to the rewards offered by a rule in which there is no penalization for wrong answers, i.e. $z=0$, and the reward for omitted answers is $y=1/M$, which coincides with the expected value of guessing. This rule, which I dubbed the *positive correction* (PC) rule, was proposed in Traub *et al.* (1969), and Budescu and Bar-Hillel (1993) point out empirical evidence on test takers preferring this rule to the SC rule. Bereby-Meyer *et al.* (2002) also found that the level of confidence required to pick an answer (among $M=5$ options) was lower under the PC rule than under the SC rule. Notice that within our calibrated model, the profitability of

[9] If $M=2$ (as in True/False questions) fairness implies that random guessing is a dominated strategy for calibrated test takers.

guessing is $w^+(p) - v(\frac{1}{M})$, which is always positive if $p \geq 1/M$ and $M > 4$ and negative if $p \leq 1/M$ and $M \leq 4$ (see Figure 3). Thus, this computations show that the PC rule is "almost" strategically neutral: for $M = 4$, for instance, random guessing is a dominated strategy for the representative calibrated test taker, but the same representative test taker will select any option with subjective probability of success $p > 0.257$ (a merely 2.8% above the level of no knowledge).

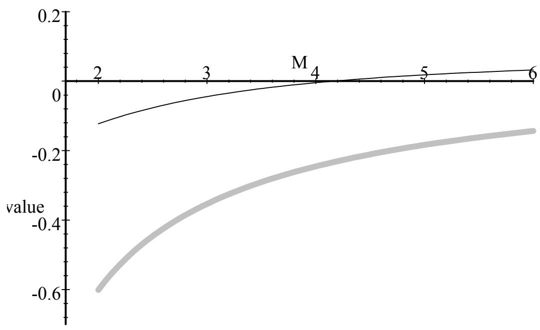

FIGURE 3: The value of pure guessing under the PC rule (black) compared with the value under the SC rule (gray).

Bereby-Meyer *et al.* suggest that loss aversion is the dominant force driving subjects' choices (as in Tversky and Kahneman's calibrated preferences) while Traub *et al.* found a higher tendency to *omit* under the PC rule. Since the profitability of guessing for the PC rule is $w^+(p) - v(\frac{1}{M})$, whereas for the SC rule is $w^+(p) + w^-(1-p)\, v\left(\frac{-1}{M-1}\right)$, one should expect more guessing under PC (as Bereby-Meyer *et al.*'s experimental data suggest) if $v(\frac{1}{M}) < -w^-(1-p)\, v\left(\frac{-1}{M-1}\right)$ for all $p \geq 1/M$, and less guessing (as Traub *et al.* suggest) if the reverse inequality holds. That would be the case if, for a given M, the judged probability of mistake, $1-p$, is small (something we expect to be associated with a high level of knowledge) and the test taker does not put too much weight on this event. Therefore, for a given level of judged knowledge, test takers overweighting small chances should behave closely to the way Bereby-Meyer *et al.* suggest, whereas test takers who do not display this possibility effect may behave closer to the way described in Traub *et al.*'s experiments.

4 Concluding remarks

A traditional argument favoring the use of the standard correction-for-guessing rule in tests has been an appeal—often implicit—to risk neutrality. The upshot of this work is that it confirms that, when we adopt a richer theory of choice, as I do here with prospect theory, the standard correction-for-guessing clearly hurts high-ability subjects by inhibiting them from using significant partial information, or "hunches" (and then guessing advantageously). Once we identify loss aversion as responsible for this situation, one way out of the problem is to award partial score credit for omitted questions rather than deduct score credit for wrong answers. The calibration presented here supports this approach showing that, for natural values of the number of options per item (usually four or more), the PC rule is close to a fair and strategically neutral rule.

At this point someone might question an item-by-item approach, and argue that a test is about a collection of repeated decision problems and not an isolated one. Be that as it may, in the interplay between single and repeated gambles it is crucial the role of framing, one of the psychological basis of prospect theory. In particular "narrow framing", the tendency for the subjects to treat each gamble as a separate event rather than integrate the series of gambles into a distribution of possible outcomes, has been not only extensively documented in experiments but it has also proven itself useful in explaining data in financial markets (Benartzi and Thaler (1995)). I can see no reason why "myopic loss aversion" should work in markets but not in a test situation. If myopic loss aversion holds, the bias against partial knowledge induced by the SC rule will not change significantly with the number of questions. This conjecture is supported by experiment 2 in Bereby-Meyer *et al.* (2002), which found no significant effect of the number of questions on the level of confidence required for answering.[10]

A perhaps more challenging problem which deserves further research is the tendency to over-diversify choice in repeated choice problems under uncertainty.[11] This effect, together with the fact that test makers show a tendency for hiding correct answers in middle positions[12] might generate distortions when going from the study of behavior in a single isolated item to answer sequences in a battery of questions.

References

Attali, M., and M. Bar-Hillel (2003): "Guess where: the position of correct answers in multiple-choice tests items as a psychometric variable," *Journal of Educational Measurement,* **40** (2), 109–128.

Benartzi, S., and R. Thaler (1995): "Myopic loss aversion and the equity premium puzzle," *Quarterly Journal of Economics,* **110**, 73-92.

Bereby-Meyer, Y., J. Meyer, and O.M. Flascher (2002): "Prospect theory analysis of guessing in multiple choice tests," *Journal of Behavioral Decision Making,* **15** (4), 313-327.

Bradley, I. (2003): "The representative bettor, bet size and prospect theory," *Economic Letters,* **78**, 409–413.

[10] The number of questions did have a significant effect, however, if the number of points gained or lost with each question was allowed to depend on the number of questions.

[11] See Rubinstein *et al.* (1997) and Rubinstein (2002) for experimental evidence collected over this topic.

[12] Attali and Bar-Hillel (2003) report empirical evidence on this topic and provide reasons for its extension to long sequences of questions.

Budescu, D., and M. Bar-Hillel (1993): "To guess or not to guess: a decision-theoretic view of formula scoring," *Journal of Educational Measurement,* **30** (4), 277–291.

Camerer, C., and T. Ho (1991): "Non-linear weighting of probabilities and violations of the independence axiom," unpublished manuscript, University of Pennsylvania

——————(1994): "Violations of the betweenness axiom and nonlinearity in probability," *Journal of Risk and Uncertainty,* **8**, 167-196.

Fox, C.R, and A. Tversky (1998): "A belief-based account of decision under uncertainty," *Management Science,* **44** (7), 879–895.

Jacobson, R.L. (1993): "Critics say Graduate Record Exam does not measure qualities needed for success and is often misused," *The Chronicle of Higher Education,* March 27-28.

Jullien, B., and B. Salanié (2000): "Estimating preferences under risk," *Journal of Political Economy,* **108**, 503–530.

Kahneman, D., and A. Tversky (1979): "Prospect Theory: An analysis of decision under risk," *Econometrica,* **47**, 263–291.

Rubinstein, A. (2002): "Irrational diversification in multiple decision problems," *European Economic Review,* **46**, 1369–1378.

Rubinstein, A., A. Tversky, and D. Heller (1997): "Naive strategies in zero-sum games." in *Understanding Strategic Interaction—Essays in Honor of Reinhardt Selten* (W. Albers, W. Güth, P. Hammerstein, B. Moldovanu, and E. van Damme, Eds.) Heidelberg-New York-Tokyo: Springer-Verlag, 394-402.

Traub, R.E., R.K. Hambleton, and D. Singh (1969): "Effects of promised reward and threatened penalty on performance of a multiple-choice vocabulary test," *Educational and Psychological Measurement,* **29**, 847–862.

Tversky, A., and D. Kahneman (1992): "Advances in prospect theory: Cumulative representation of uncertainty," *Journal of Risk and Uncertainty,* **5**, 297-323.

Tversky, A., and P. Wakker (1995): "Risk attitudes and decision weights," *Econometrica,* **63**, 1255–1280.

Wakker, P. (2001): "Testing and characterizing properties of nonadditive measures through violations of the sure thing principle," *Econometrica,* **69**, 1039–1059.

Wu, G, and R. González (1999): "Non-linear decision weights in choice under uncertainty," *Management Science,* **45** (1), 74–85.