Noesis

Editorial

Ronald K. Hoeflin
P. O. Box 7430
New York, NY 10116

Dues: I have received dues payments from four members since the announcement of the new $20 annual fee: from Richard May, H. W. Corley, Ray Wise, and Dean Inada. Since I announced that dues would be reduced by one dollar for each page of material submitted, I have also received six pages of material from Eric Hart and six pages from C. M. Langan. Eric's submissions were included in Noesis 24 and Mr. Langan's will occupy much of this issue.

Comments on Langan's remarks: Mr. Langan makes three specific criticisms of past material in Noesis: (1) concerning problem 26 in Trial Test "B", which appeared in issue 14; (2) concerning Chris Cole's solution to Newcomb's paradox in issue 15; and (3) concerning the new norming method by Keith Raniere for the Mega Test, which I tried to outline in issue 23.

As for (1), several prople, including Chris Cole and Dean Inada, drew my attention to the shortcomings in my formulation of the marble problem, referring me to the Bayes theory of probability. I had never heard of Bayes, but according to my New Columbia Encyclopedia he was an English clergyman and mathematician who lived from 1702 to 1761, and his "Essay Towards Solving a Problem in the Doctrine of Chances," first published in 1763, sought to establish that "the rule for determining the probability of an event is the same whether or not anything is known antecedently to any trials or observations concerning the event." In any case, I did propose alternative formulations of the marble problem in a subsequent trial test.

As for (2), I believe that Dean Inada also disagreed with Chris Cole's solution to Newcomb's paradox, but Dean did not go into detail. The only Newcomb mentioned in my Columbia is a Simon Newcomb, who was an American astronomer and mathematician who lived from 1835 to 1909. But the encyclopedia makes no mention of a paradox named after this Newcomb.

As for (3), I have asked Keith to write up a clear and simple booklet explaining his norming method, which I only partially understand. I think this norming method, if it is viable, would make an interesting and worthwhile topic for a doctoral dissertation in the field of psychometrics. Of the 3,071 doctorates awarded in the field of psychology in the U.S. in 1986, only 11 were in the area of psychometrics, which may be an undercultivated subdiscipline, judging by this low number.

Martin Gardner's letter: I conclude with a copy of a letter from the famous Scientific American game columnist of yore, Martin Gardner, to member Richard May, who wrote to Gardner concerning a series of board games of a geometrical nature that he designed.

Two Letters to the Editor

C. M. Langan
P.O. Box 131
Speonk, NY 11972

Dear Mr. Hoeflin:

Sorry if I seemed to ignore your announcement of Mr. Cole's visit to New York. Unfortunately, I'm not within easy range of the city and I was short on time that day. I'm originally from California myself, having been born in San Francisco. Though I lived briefly in the Los Angeles/Hollywood area as a young child, my recollections don't go much beyond an endless rank of swimming pools, one per back yard, and sometimes awakening at night with great difficulty in breathing. I gather things haven't improved that much with regard to air quality.

I'd have sent an introductory letter several months ago, but the first draft seemed to lack the right amount of friendliness. Since I had no desire to appear unsociable, and since it ran to over a page in length, I simply let it hang. Now that I understand you to be entertaining longer pieces, and with the assurance that I intend to tax nobody's pride or credulity, it went partly as follows:

'Thanks for sending the back issues of Noesis, in which I found a couple of questionable entries. Since these concern important unresolved issues in mathematics and philosophy, it may be inadvisable to let them stand unchallenged.

'The first of these entries (Issue #14) concerns problem #26 on your "Trial Test B". It is posited that the solution depends on the number of colors among which the marbles are distributed, and further, that the absence of such information entitles us to certain suppositions concerning the distribution. This may be intuitively plausible, but actually amounts to an apparent misuse of the so-called "principle of indifference".

'Concisely, while the probability of the sample data may be a function of the distribution of marbles among a certain range of colors, the same thing may be said in reverse: the hypothetical range and distribution of colors is probabilistically dependent on the sample data. According to the wording of the problem, the interplay between sample and distribution, by which any independent assumption regarding one induces dependency in the other, is constrained by the fact that the outcome of the sampling procedure is given in advance. That is, it is certain, a constant, and thus independent. The distributional hypothesis is what is dependent, and any "assumption" regarding it must be inferred from the sample data. Since this is a contradiction, there can be neither such an assumption nor a need for one in this case. While the distribution obviously precedes the sample in real time, this becomes irrelevant with respect to dependency relations between variables.

'The remarks made about this problem evoke the wider context of confirmation theory, the concern of which is the evidential and implicative verification of hypotheses. Its major open debates center on the logical foundations of the theory of probability.

- 2 -

'The second arguable contribution (Issue #15) concerns a dilemma known as Newcomb's problem, which shares the consideration of dependency. The correct solution of this "paradox" - and it is not so easy to find the real logical paradox within the problem - might be expected to involve some subtle, powerful reasoning. As it turns out, the "argument from temporal dominance", on which the proposed solution is expressly based, is demonstrably inadequate to address the matter in its full complexity. So, as it happens, is the countervailing "argument from expected utility", which is a simple application of elementary decision theory and delivers the other possible answer. For reasons that go beyond either argument, the subject of the prediction should take only box B.

'This problem is of particular interest due to its bearing on the nature of time and the question of free will, connections which remain somewhat obscure despite several published analyses. The ramifications are extensive and include many fertile topics.'

Since this was written, I've become aware of several other variously explicit assumptions with which I'm afraid I must differ. One of these apparently involves my delayed invitation to join this group: that the Mega test retains its supposed ability to discriminate among the extremely intelligent when administered in conjunction with a set of absolute cut-off scores which imply that the test has already been normed.

There are several mutually-reinforcing arguments, based on statistics, decision theory, and computer science, to the opposite effect. In order, they involve statistical uncertainty, the equation of maximum gain and maximum advisable expenditure in certain types of games, and the nonequivalence of problems with respect to the amount of time needed to solve them. They add up, with mathematical precision, to an irrefutable conclusion: the best way to take a pre-normed, gain-maximized Mega-style test is to solve the number of problems sufficient for maximum gain roughly in order of ascending difficulty, and when certain of having done so, either stopping or guessing on the remainder to minimize expenditure of time and effort. To do otherwise, while perhaps of recreational value for those addicted to compactly-expressible puzzles, is clearly inefficient for those with little time to spare, which description covers most highly intelligent people. Concisely, it is to lower the ratio of gain to expense, or that which any rational strategy seeks to maximize. This argument - and it may be considerably embellished - suffices to show that there is no reasonable counterargument by which those passing the test under the above conditions could, by its means alone, be held less intelligent than anyone else taking the test, regardless of scores achieved.

Of course, higher-scoring pre- or post-Omni participants may not have encountered the same misleading information. If not, their reward still cannot rationally exceed that of the others, who in effect were unfairly handicapped by the injudicious

publication of absolute cut-offs. I see no way around this, and in
my opinion, it should be spelled out in view of the internal
stratification which which seems to have elevated members of the
new "Mega Society" above those "Titans" meeting the former
explicit qualifications. To put it bluntly, not to remedy this
blunder is to let unscientific procedure undermine both the
definition and goals of the society, whatever they may be.

I almost hesitate to mention the recently-proposed method for
renorming the Mega test, since I'm unsure of the accuracy or
completeness of the description offered. However, it had better be
pursued carefully if at all, and for the following reasons.

There is an obvious source of error in this procedure, which
we may define tentatively as the average discrepancy between the
I.Q. implied by the hardest problem solved by an examinee, and his
overall score. To explain, I present one extreme example. Say an
examinee solves problem #36 on the Mega test, but no other. Then
he can either be given a placement based on the solution-
frequency of #36, or he can be given a placement based on the fact
that he solved only a tiny fraction (1/48) of the problems posed,
or he can be placed according to some compromise between these
factors. If his placement depends on anything but his overall
score, however, then this is to admit the full range of conjecture
as to why he missed all the easier problems. We are driven, by the
relative infrequency of the one correct solution, to a
consideration of *motivation* - one reason the examinee solved only
#36 may be that it "interested" him more than the others. But this
possibility, and any other of like effect, is destructive to basic
assumptions in the design of the test as a measuring device: it is
intended to measure not motivation, but *intelligence* (an
interesting distinction in itself, and highly suspect). This can
introduce a huge uncertainty in the calibration. The only way to
suppress this uncertainty is to exorcise the source of error,
which reduces to the ordering of problems by difficulty. But this
order is what the proposed method relies on, however indirectly,
and to narrow the percentiles under these circumstances is a pure
negative circularity.

In other words, any variability in the test data with respect
to the specific problems solved by participants at or below any
given stratum strongly reduces the relative bias of the problems
involved, which in turn limits the quantisation of strata. If we
rely instead on just those problems *consistently* exhibiting high
bias, we reduce allowable variability, invalidate a large
percentage of the overall scores, and are left only with the
solution-frequencies of a few problems relative to measurements on
other tests. This, by the way, reduces to a well-known generality
in statistics: high bias tends to accompany low variability, and
vice versa. The "bias" of a problem, of course, is the extent to
which it discriminates by intelligence; analogous to the asymmetry
of a loaded die, the bias (or difficulty) of a problem is what
tends to make it come up unsolved more often than solved.

If, for that matter, the data display little variability - if
we find high regularity in the specific problems solved by those
achieving the same overall scores - then we can still not do much
better than use the absolute solution-frequencies to learn
wherein resides the bias, use this knowledge to create an extended
sequence of increasingly difficult problems, and base each score
on the hardest one solved in each of several disparate categories.
This method, which still contains some difficulties, would at
least spare an examinee the onus of leaving markers along the
scale, requiring only that he present himself at the end of his
intellectual rope. Notice that this is precisely how the world
evaluates genius, bestowing the title on those solving as few as
one important but extremely difficult problem.

Enough concerning questionable assumptions; the rest can
wait. Irrespective of my belief that it cannot successfully be
done, other members are free to disagree with any of the above
assertions. However, they should take care in the process to make
adequate distinctions between matters of opinion and logic. I
don't, of course, rule out the possibility that others, including
those with whom they originated, have spotted the supposed errors
but elected to remain silent. Were I to become convinced that the
journal can accomodate anything more than brief and inadequate
summaries of the full developments, I'd consider providing more
exact reasons than now seem appropriate. My present intention is
only to relieve the membership of the collective embarrassment of
claiming intellectual superiority while breeding canards, all of
which sooner or later end up plucked.

Formal education, at least as currently structured, has very
little to do with high intelligence or creative ability. I'll
therefore prune my academic history to the mention of a brief
stint at Reed College, known at one time as a mecca for the
intellectual nonconformist. There were too many distractions and
too little classroom motivation, and the course of my studies -
which nominally centered on mathematics and philosophy - tended
to be idiosyncratic. The distractions included a few memorable
visitors, more or less exemplary of whom were the novelist Ken
Kesey and his magic bus of merry pranksters. I remember the place
as being lunatic, but that description would seem to fit most
colleges and universities to some extent.

It appears that I may be occupied on Independence Day weekend,
though I don't yet know for certain. Again, it is not my wish to
jeopardize the cooperative aspects of this undertaking, but only
to preclude such action from outside. If not for that, I probably
wouldn't bother to take exception.   C. M. Langan

Dear Mr. Hoeflin:

I've just received your acknowledgement of my recent communication. On reading it over, I see why you might think it "still seems highly negative". This, as I tried to maintain, was not my intention. It's difficult sometimes to take exception to mathematical hypotheses without seeming overly aggressive about it - I'm sure you've seen a few papers in which the spirit of intellectual competition was inadvertently overdone. I was not, as you seem to imply, trying to bring you or anyone else to account for the disputed entries, which plainly were made in a casual way. I'm sorry if you found the tone distressing.

You're only partly responsible for the matter of cut-offs. I agree with your opinion that it's better to be safe than sorry when taking such a test, but must question your doubt concerning anyone's ability to "fine-tune their efforts" to be certain of attaining a given score. Assuming that you have a reasonable amount of faith in the unique correctness of your own answers, it should be obvious to anyone intelligent enough to score near your test's ceiling when he is guessing (uncertain) and when he has found, implemented, and checked a definite recursive procedure for solution. It may be true that *most* people are incapable of seeing the inadequacies in whatever their recursive lines of attack on extremely difficult problems. But as intelligence rises towards the ceiling of the test, these shortcomings clearly become more obvious. As you might recall, I did give 44 correct answers, but two of them after what I considered to be better, though similar, first choices (I don't see how one can make sure of the absolute preferability of specific answers to some verbal analogies given that natural language has ambiguous and relativistic aspects; perhaps more than one answer could be allowed for variant analogies without compromising their efficacy). Also, on looking the test over again, I see that the other problems I missed were just those on which I intentionally guessed. My reasons for guessing on one had to do with a boring chain of calculations I felt like trying to circumvent, and for a couple of others, there is as yet no quick and easily-stated method for generating solutions. I notice that some of your "spatial" problems have this quality, all but forcing one either to start exhausting possibilities, or to make leaps of intuition for which there is seldom an obvious need. Such problems are basically optimizations of the kind usually left to indefatigable computers that don't mind eliminating numerous false optima before tracking down the real one.

These aren't excuses, of course, but I'm confident that I could have improved my score with just a little more effort, even allowing for the two verbal problems on which we differed. I had no invitation to try again in view of my initial qualification. Aside from that, however, I understand the difficulty of creating, administering, and evaluating a test like yours, and congratulate you on a fine effort.

There might also be a little confusion concerning the other

two matters for which you offered to stand. I see nothing the
matter with your marble problem, and am simply trying to
straighten out a bit of circular reasoning concerning it. I don't
even see anything particularly heinous about tossing out a little
circular reasoning occasionally, since problems that don't get
stated, even in a roundabout way, often don't get solved as soon
as they might. There's a difference between being wrong
and clinging obstinately and without higher purpose to incorrect
reasoning even when you know you can improve on it, and I doubt
thoroughly that any Titan Society members would favor such a
strategy for its own sake. Every intellect, no matter how careful,
sometimes lets a mistake slip by, and it doesn't make him or her
any less intelligent.

Nor do I mean to imply that what you call the Raniere norming
method may not contain promising insight. I only wanted to point
out that, according to mainline statistical theory, there is an
uncertainty principle that imposes limits on what can be inferred
by statistical means alone. In fact, I'm waiting for a fuller
explanation. Even if such a method can't be used to narrow
percentiles past a certain bound, it might still help confirm them
all the way up to it.

I suppose I ought to mention Newcomb's problem as well, just
to be thorough. Many intelligent people see time as dominating
reality; once something has happened, it has happened, period. But
the context of this problem is not restricted to the point at
which the outcome is irrevocable. It is defined on a set of
conditions which refer to a whole range of temporal predicates.
Those who have argued purely from game theory have never been
granted a unanimous victory over the problem precisely because
they were unable to deal with the multitude of its aspects that go
beyond mathematical expectation. It was first posed by a physicist
(William Newcomb), and physicists don't as a rule consider
themselves obligated to accept game-theoretic arguments, which can
seem devoid of physical insight, as decisive. I think I can do
better only because I happen to have worked on the problem in
conjunction with a related theory.

Once again, I'm usually a very friendly person, and don't
want anyone to think otherwise. People attempt to correct each
other all the time, and only those with the thinnest skins believe
themselves entirely immune. The intelligence of the membership,
and of you, was never in serious doubt here, nor would I presume
that mine was there.

Don't hesitate to print this with my other letter,
particularly if you feel that this would ameliorate its draconian
appearance, or to let me know if you consider it too harsh
entirely. If so, I could try to trim off some of the rough edges.
Regarding which, you can forget the remark about "canards" - I was
just having a little fun, probably a big factor in any group like
this one. Enough so to be excusable, anyway.

My best to the membership.   Christopher Langan

- 7 -

New address

Martin Gardner, Pres.   /110 Glenbrook Dri...
Woods End, Inc.
~~103 Woods End Drive~~
Hendersonville, North Carolina 28739

Dear Mr. May: 20 Mar 88

I honestly can't tell you whether I received Tri-hex
or not. For some reason, I receive lots of board games
from inventors, but since I don't consider myself an expert
on such games, I routinely pass them along to my friend
Sid Sackson (1287 Arnow Ave, Bronx, NY 10469). Sid is a
board game inventor, and owner of the largest private collection
of math board games. His opinion is worth a hundred times
more than mine. He also evaluates board games for a periodical
devoted to such things (I forget the name).

All best,

Martin Gardner

A Final Word from the Editor

For those contemplating coming to the meeting of this group
on the July 4 weekend, I should mention that if you want to save
a lot of money on transportation into the city from the airports,
take the Q10 bus from Kennedy Airport to the subway and then take
the E train to 8th Avenue and 50th Street, or take the Q33 bus from
LaGuardia Airport to the subway and then take the E train (again)
to 8th Avenue and 50th Street. These buses stop right across the
street from the airport terminals. The bus fare is one dollar
(you need exact change, not folding money), and the subway fare is
another dollar, for a grand total of two dollars. A taxi would cost
you ten or fifteen times as much ($20 or $30). If you arrive at
Newark Airport, you can get a bus into New York City for five dollars,
which stops at the Port Authority Bus Terminal at 8th Avenue and 42nd
Street. You might have to ask someone where this bus stops. If
you prefer, I will meet you at any of these airports so long as
you don't ask me to split the cost of a taxi into the city, since
I prefer the less expensive bus-and-subway routes.

- 8 -