

## **A CRITIQUE AND REINTERPRETATION OF GORDON'S IQ - COMMENSURABILITY PROPERTY**

by J. Michael Bailey, Department of Psychology, The University of Texas, Austin, Texas 78712.

In an earlier issue of this journal (Vol.7, 1987, pp. 30-96) Gordon presented a model based on IQ which, he claimed, explains black-white differences in rates of juvenile delinquency. The explanations involved the demonstration that prevalence rates of juvenile delinquency satisfy a property which Gordon termed "commensurability" with IQ (IQ-commensurability). Furthermore, he argued that similar models based on SES or education are unsatisfactory because these variables fail the test of commensurability with respect to delinquency rates. Largely on the basis of these analyses, he concluded that the black-white difference in IQ is much more credible than the difference in either education or SES as an important cause of race differences in delinquency. The purpose here is to show that the property of IQ-commensurability is largely irrelevant to any reasonable explanation of the race difference in delinquency rates. Furthermore, IQ-commensurability can be partially explained by an existing model which is highly plausible, has demonstrated construct validity, and can be viewed as a logical extension of Gordon's own model.

### **IQ-Commensurability and Race-IQ-Delinquency Models.**

*Background.* Whites differ, on average, from blacks in IQ (with whites' scores higher, on average) and in rates of delinquency (with blacks' rates being higher). Whether and how much race differences in IQ contribute to race differences in delinquency is a reasonable, if controversial, question (Wilson & Herrnstein, 1985), and was the principal topic of the aforementioned paper by Gordon.

Gordon examined two independent sets of delinquency rates. One set pertained to the criterion of obtaining a juvenile court record in Philadelphia between 1949 and 1954. The other set included the rates of commitment to a U.S. Training School in 1964. Each set included separate rates for blacks and whites by sex. The Training School rates reflected an apparently much more severe criterion for delinquency. For instance, the percentage of young white males in Philadelphia who obtained a juvenile court record was 17.86%. The percentage of young U.S. white males who were committed to Training Schools was 1.01%.

Using these two quite different sets of delinquency rates, Gordon provided a remarkable demonstration. For each prevalence rate he

computed a "critical IQ" below which an identical proportion of individuals of a given race is found. For example, .5086 is the proportion of black males meeting the Philadelphia criterion. This corresponds to a critical IQ of 86.3, because the proportion of blacks obtaining an IQ below 86.3 is .5086. The 17.86% of white males who met the same criterion corresponded to a critical IQ of 86.7, because the proportion of whites obtaining an IQ below 86.7 is .1786. The critical IQs for black and white males were very similar, though the actual respective delinquency rates were quite disparate. Likewise, the critical IQs were similar for black and white females who met the Philadelphia criterion: 73.5 and 71.8, respectively, for the rates of .1582 and .0335. Critical IQs also matched well for the U.S. Training School Data. The prevalence rates for black and white males were .0400 and .0101, respectively, which yielded critical IQs of 60.9 and 63.7. For black and white females, the respective rates were .0082 and .0023, yielding critical IQs of 52.4 and 55.3. For each data set, then, the critical IQ computed for black males was quite similar to that computed for white males. The critical IQ's of black and white females were also near each other, though they differed substantially from those of males. Gordon referred to this fact of matching critical IQ's as "IQ-commensurability".

*Irrelevance of IQ-commensurability: An overview.* What, if anything, does IQ-commensurability imply about the nature of race differences in delinquency? More precisely, does the fact of IQ-commensurability provide evidence for a plausible model of black-white IQ differences causing black-white delinquency differences? Contrary to Gordon, I shall argue that it does not.

The argument will proceed as follows: (1) The model which is directly suggested by IQ-commensurability is clearly false. (2) A second, more plausible model considered by Gordon is both empirically false and inconsistent with IQ-commensurability. (3) The remaining model which Gordon considers is so general that it may be true; however it is consistent with a wide range of race differences in delinquency, most of which fail to generate IQ-commensurability. (4) There is no known theory of race differences in criminality which would predict IQ-commensurability, as such.

*Race-IQ-Delinquency Models.* The model which is directly suggested by IQ-commensurability requires that any individual whose IQ is below the (gender-appropriate) critical IQ be delinquent; no individual whose IQ exceeds the critical IQ can be delinquent. Gordon reasonably rejects such a model as "not at all plausible" (p.35). However, he desig-

nates it a "proxy" model which "should be sensitive to the correctness of the model that is really assumed" (p.35).

In order to evaluate Gordon's claims for the model which is assumed, it is necessary to specify the latter. The first reasonable model which Gordon mentions is that blacks and whites have identical IQ-specific delinquency rates. This would imply that the black-white difference in delinquency is due solely to an excess of blacks at (low) IQ-levels which are most likely to lead to delinquency. Gordon has tested this model via computer simulations, and has found that it is false. The nature of the failure is that the model underpredicts delinquency rates for blacks whenever, as in the present case, IQ-commensurability obtains (and plausible IQ-specific delinquency rates are used). The magnitude of underprediction varies over the IQ range, but is on the whole substantial. Thus, IQ-commensurability does not support a theory of race differences in delinquency based solely on identical IQ-specific delinquency rates.

The second plausible model that is considered, and which is apparently the one eventually "really assumed", attempts to account for the residual black delinquency unexplained by the first. According to this model, race differences in delinquency rates are a function not only of the proportionate excess of blacks at IQ levels predisposing them to delinquency, but also of a number of other statistical and sociological factors. These factors include the lower average IQ of black parents, which is hypothesised to lead to poorer socialisation; greater density of delinquency within the black population causing feedback effects; and other "contextual effects" - presumably including those of a more traditional sociological nature, such as the excess of broken homes among blacks.

The important point to note here is that the set of variables thought to cause the black excess in delinquency (beyond that explained by IQ) is not completely specified. A corollary is that neither the magnitude nor the distribution of their effects is specified. Without such specification, the predictive power of the model is quite weak, as it is consistent with a wide range of values for race differences in delinquency. In fact, it is consistent with any values of race differences in delinquency larger than those predicted from the IQ differences alone.

Gordon admits that "there is no simple mathematical relation between the property of IQ-commensurability and IQ-specific delinquency rates". This is essentially an admission that there is no strong theory (even post hoc) as to why the effects of IQ and the additional variables should sum to produce the property of IQ-commensurability.

It is evident that the property of IQ-commensurability provides very little support for the model of race differences in delinquency caused by differences in IQ when other variables are added. By the same token, the failure of education and SES to be commensurate with delinquency prevalence rates is not evidence that these variables are less important than IQ in causing race differences in delinquency. If we allow an analogously vague model for education or SES, there is an endless number of delinquency prevalence rates (including the true rates) which would be consistent with the model.

To be sure, Gordon is correct in pointing out certain advantages for IQ over education or SES in explaining race differences in delinquency. For instance, race differences in IQ are larger than those for education or SES and so, have more potential explanatory power; race differences in delinquency have remained fairly constant over the past few decades, like IQ, but unlike SES or education; in explanatory models, SES and education are often proxies for IQ. But these facts are known quite independently of the fact of IQ-commensurability. Gordon's thesis regarding the relative importance of IQ as a factor in race differences in delinquency may well be true; however IQ-commensurability adds little or no support to his argument.

### **An Alternative: the Multifactorial Threshold Model of Delinquency.**

*The model.* Though it fails to provide support for Gordon's model, IQ-commensurability remains an interesting phenomenon in search of an explanation. Fortunately, there is a model, the *multifactorial-threshold model*, which may help. The model has at least two advantages: (1) It can incorporate Gordon's most reasonable model, and (2) it has demonstrated construct validity for a phenomenon which is quite similar to delinquency. In Gordon's model, recall, individual differences in delinquency--both between and within races--are caused by IQ and an (incompletely specified) array of other factors. This means that delinquency is "multifactorial", in the sense of having many causes.

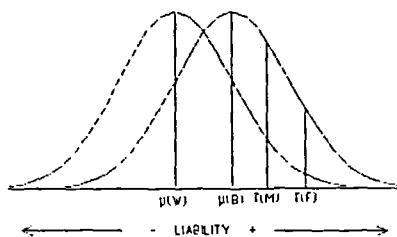
If we consider the sum of all the effects, it is reasonable to hypothesise that the resulting distribution will be approximately normal. This would be the case if the factors besides IQ are both numerous and, on average, only moderately intercorrelated. If so, the distribution of their summed effects,  $Y$ , will be approximately normal. Since IQ is normally distributed, the sum,  $IQ + Y$ , will also be near normal. This distribution of the summed effects of causes of delinquency is called the *liability* to delinquency.

The *threshold* of the model is that point along the liability distribu-

tion beyond which we call individuals "delinquent". The threshold is required by the fact that most definitions of delinquency—including those in the data sets examined by Gordon—are qualitative. Obviously, different definitions of delinquency will yield different thresholds. Definitions reflecting more severe criteria will require thresholds further from the mean of the liability distribution. In the present case, for instance, commitment to a state school represents a more severe form of delinquency than acquisition of a juvenile court record.

The multifactorial threshold model was first used in medical genetics to conceptualise disorders such as diabetes (Falconer, 1965), but has become increasingly common in psychiatry (e.g., Gottesman & Shields, 1982). Cloninger and Gottesman (1987) have provided evidence for the model's construct validity for antisocial personality disorders (ASPD), a condition which is highly relevant to juvenile delinquency. There is a large sex difference in rates of ASPD, the disorder being much more common among men. There is also a sex difference in familiarity patterns. A female with the disorder has, on average, significantly more affected relatives than does an affected male. Cloninger and Gottesman hypothesised that ASPD is a multifactorial char-

Fig. 1: Multifactorial Threshold Model of Delinquency with Separate Distributions for Blacks and Whites, and Separate Thresholds for Males and Females



$\mu(W)$  = the mean liability of whites  
 $\mu(B)$  = the mean liability of blacks  
 $T(M)$  = the delinquency threshold for males  
 $T(F)$  = the delinquency threshold for females

acteristic with two thresholds: one for males, and a separate one for females, the latter of such a nature that women need more of the liability to have the disorder. Using this model, strikingly similar values were computed for both sexes for the familiarity of the underlying liability.

Although the two phenomena are related, ASPD is restricted to severe and chronic patterns of criminality (American Psychiatric Association, 1987), not necessarily in minors. Given the similarity of the relevant behaviours, however, the causes of ASPD likely overlap considerably with those of delinquency. Thus, the success of the multifactorial threshold model in representing ASPD supports the plausibility of the model for delinquency.

*The model applied to prevalence rates.* Figure 1 depicts a multifactorial threshold model of delinquency which represents both race and sex differences in prevalence rates. What were separate black and white IQ distributions in Gordon's IQ-commensurability are now separate black and white distributions of liability. (The figure has been oriented so that liability increases from left to right.) The black liability distribution is to the right of the white distribution, indicating that, on average, blacks have a greater liability to delinquency than do whites. What were critical IQ's in Gordon's demonstration are now threshold values. Following Cloninger and Gottesman, there are two thresholds in Figure 1: one for males, and one for females. Figure 1 thus implies that for each sex, the same threshold applies to both races. The fact that the female threshold is to the right of the male threshold means that females require a more severe "dosage" of liability factors than males do in order to become delinquent, thus there will be fewer females than males meeting a given delinquency criterion. An explicit assumption of the model is that the same factors contribute to the liability to delinquency for both males and females.

There are two parameters which, together with the normal shape of the distributions, provide all information about the model: (1) the ratio of the black to the white standard deviation, and (2) the difference in the means of the two distributions. Gordon showed that the respective IQ parameters do a good job of representing the delinquency data in such a model. Gordon gave no convincing rationale for the model, itself, but instead stressed the alleged meaningfulness of the fact that delinquency parameters are commensurate with IQ parameters. I have shown that we have no reason to believe that IQ-commensurability implies anything about the importance of IQ as a cause of race differences in delinquency. The similarity between the black-white differences in

delinquency and IQ cannot be explained by the model under consideration as anything but a fortuitous coincidence.

Lest one should think IQ-commensurability too striking to be a mere coincidence, it should be noted that not just the IQ parameters are commensurate with the prevalence rates examined by Gordon. Several possible sets of values yield a good fit, defined as an average of at least 98% reduction in between group variance (see Gordon, 1987, p. 43.) over both sexes and both data sets. These values range from .80 to 1.20 (white standard deviations) for the black-white mean difference, and 0.60 to 0.85 for the black-to-white ratio of standard deviations. It is somewhat less difficult to accept that IQ-commensurability is a mere coincidence upon the realisation that IQ parameters are but among a range of values which yield a good fit to the model. (In order to determine which--if any--is the "true" set of parameters, more data points are needed.)

Table 1 lists threshold values for four sets of parameters which yielded very close fits to the data. The thresholds were computed as z-scores relative to the mean of the white distribution, which was set at zero. The thresholds were computed so that higher numbers imply greater liability. The nearer the thresholds of blacks to those of whites (within sex and data set), the better the fit. The following example demonstrates the interpretation of threshold values: if we assume that the black-white difference in liability to delinquency is 0.90, and the black-to-white ratio of standard deviations is 0.80, then black males in Philadelphia between 1949 and 1954 with a liability to delinquency greater than 0.88 standard units acquired a juvenile court record.

Strictly speaking, the data concerning Philadelphia court records and the data concerning state school commitment should not be considered simultaneously. This is because the liability distribution of the former is most accurately interpreted as "liability to acquiring a juvenile court record in Philadelphia between 1949 and 1954". The proper interpretation of the latter's liability distribution is analogously restrictive. Of course, the two liabilities should be closely related. To see this, it is useful to conceptualise each liability as composed of one part which both share and one part which is unique. The shared part might be conceived of as "liability to sanctions for criminal behaviour." The part unique to the state school commitment liability would include, for example, factors present in 1964 but not during 1949-54, national as distinct from Philadelphia factors, etc.

*Implications of the model.* Despite the limitations imposed by this caveat, there are several interesting consequences when Gordon's dis-

**Table 1: Four Sets of Parameters Fitting Observed Prevalence Rates Under the Assumption that Delinquency is a Multifactorial Threshold Characteristic**

Parameters		Thresholds for Philadelphia Data <sup>a</sup>				Thresholds for National Data <sup>b</sup>				Mean Fit <sup>c</sup>
Mean Difference (B-W)	Ratio of SD's (B/W)	Males		Females		Males		Females		
		Blacks	Whites	Blacks	Whites	Blacks	Whites	Blacks	Whites	
0.90	0.80	0.88	0.92	1.75	1.83	2.39	2.32	2.94	2.85	99
1.00	0.75	0.99	0.92	1.75	1.83	2.31	2.32	2.68	2.85	99
1.10	0.70	1.09	0.92	1.80	1.83	2.33	2.32	2.78	2.85	99
1.20	0.65	1.19	0.92	1.85	1.83	2.34	2.32	2.76	2.85	99

<sup>a</sup>Criterion is acquisition of juvenile court record in Philadelphia between 1949 and 1954. The delinquency rates are as follows: black males, 50.86%; white males, 17.86%; black females, 15.82%; white females, 3.35% (Gordon & Gleser, 1974).

<sup>b</sup>Criterion is commitment to a US Training School in 1964. The delinquency rates are as follows: black males, 4.00%; white males, 1.01%; black females, 0.82%; white females, 0.23% (Gordon, 1973).

<sup>c</sup>Percent reduction in between group variance (Gordon, 1987).

Note. Thresholds are computed as z-values using the white standard deviation and setting the white mean at zero.



coveries are reinterpreted in light of the multifactorial threshold model. For instance, despite the non-equivalence of the two liabilities, Table 1 suggests that one set of parameters can do a good job of representing the black- white difference in each. This suggests that the unique parts of the liabilities show negligible race differences. Otherwise, one would need to accept the unlikely proposition that race differences in factors unique to one data set are equal in magnitude to those unique to the other. It also suggests that a good estimate of the race difference in delinquency is between .90 and 1.20 (white) standard deviations. If so, this gives the criminologist an estimate that does not fluctuate according to the severity of the criterion, in contrast to ratios of percentages. The size of the race difference in delinquency can now be compared to sizes of other race differences of continuous traits, such as IQ.

Because the parameters of the model are currently consistent with an uncomfortably large range of values, examination of data sets using other criteria for delinquency would be desirable in order to determine whether the range can be narrowed. It would be interesting to know, for instance, whether prevalence rates for Antisocial Personality Disorder can be fitted by the model with one of the sets of parameters. For instance, Cloninger et al. (1975) reported that the observed prevalence of ASPD in the general white population was 3.3% for males and 0.9% for females. If, for example, the black-white difference in liability to delinquency is 0.90, and the black-to-white ratio of standard deviations is 0.80, we would expect the black rates of ASPD to be about 12% for males and 3% for females.

As noted above, the model does not, by itself, point to the importance of any factor, including IQ, in causing race differences in delinquency. In fact, the model clarifies the limits of IQ as an explanatory factor of these differences. The within-race correlations of IQ with delinquency or criminality cited by Gordon ranged from -.20 to -.35. Even if we accept the larger estimate (in absolute value) and any set of parameters in Table 1, removing the race differences in IQ would leave a difference of over .5 of a standard deviation in delinquency to be explained.

An important inference can be drawn from the multifactorial threshold model regarding the nature of sex differences in delinquency and their relationship to race. The model represented in Table 1 accounts for race differences equally well for males and females-i.e., thresholds for black and white males tend to be similar, and thresholds for black and white females tend to be similar. Thus, the race differences in delinquency are approximately the same for males and females. This suggests that the causes of race differences in delinquency are similar for

both sexes. If so, theories of race differences which emphasise factors unique to one sex (e.g., the detrimental effect on black males of the lack of appropriate male role models) have limited explanatory power.

Conversely, for both sets of prevalence values, the sex difference for blacks is similar to that for whites. This implies that the causes of sex differences in delinquency are similar for blacks and whites. However, the size of the sex difference varies between the two data sets. Across all four sets of parameters in Table 1, the sex difference in the Philadelphia data is appreciably larger than that for the state school commitment data. This discrepancy has a number of possible interpretations. For instance, sex differences in delinquency may have genuinely decreased from 1954 to 1964. This discrepancy could also occur if commitment to a state school represented a less severe delinquency for females than for males. It is noteworthy that the discrepancy between the values of sex differences in the two data sets is approximately equal for blacks and whites, providing additional evidence for the independence of sex and race differences.

These implications of the multifactorial threshold model of delinquency await replication before they can be considered reliable. Nevertheless, they demonstrate possible useful applications of the model.

*Final Remarks.* Elsewhere, Gordon (1980) has written eloquently of the difficulties and iniquities which are too often confronted in the course of research on race differences. These problems are greatly magnified when one attempts to investigate differences of social importance, of which both delinquency and IQ are examples. Too often *ad hominem* attacks substitute for legitimate scientific disagreement. I hope it is clear that the present paper is not an example of the former. I concur with Gordon that it is important to study differences of social importance, and I am pleased that capable social scientists like Gordon are doing so. Although my criticisms of the concept of "IQ-commensurability" are of a fundamental nature, they are intended constructively to further progress in this area. Furthermore, I believe that they have succeeded by putting Gordon's important discoveries into a more interpretable framework.

### Acknowledgements.

The author is indebted to John Loehlin for his thoughtful comments on various drafts of this project, and to Lee Willerman for his editorial assistance.

### References.

American Psychiatric Association

1987 *Diagnostic and Statistical Manual of Mental Disorders*, Third Edition, Revised. Washington D.C.: American Psychiatric Association.

Cloninger, C. Robert, and Irving I. Gottesman

1987 "Genetic and environmental factors in antisocial behavior disorders." Pp. 92-109 in S. Mednick, T. Moffit, S. Stack (eds.) *The causes of Crime* New York: Cambridge University Press.

Cloninger, C. Robert, T. Reich, and S. B. Guze

1975 "The multifactorial model of disease transmission: II. Sex differences in the familial transmission of sociopathy (antisocial personality)." *British Journal of Psychiatry* Vol. 127: pp. 11-22.

Falconer, D.S.

1965 "The inheritance of liability to certain diseases estimated from the incidence among relatives. *Annals of Human Genetics*, Vol.29 pp. 51-76.

Gordon, Robert A.

1980 "Research on IQ, race, and delinquency: Taboo or not taboo?" Pp. 37-66 in Edward Sagarin (ed.), *Taboos in Criminology* Beverly Hills: Sage.

1987 "SES versus IQ in the race-IQ-delinquency model." *International Journal of Sociology and Social Policy* Vol.7: pp. 30-96.

Gottesman, Irving I. and James Shields

1982 *Schizophrenia: the Epigenetic Puzzle*. New York: Cambridge University Press.

Wilson, James Q. and Richard J. Herrnstein

1985 *Crime and Human Nature*. New York: Simon and Schuster.