

1 AFFIDAVIT OF STEPHEN GREENSPAN, PhD

2 Personally appeared before me the undersigned officer
3 duly authorized by law to administer oaths Stephen Greenspan,
4 who being duly sworn or affirmed, states as follows:

5 My name is Stephen Greenspan. I am of legal age and
6 competent to testify as to matters set forth in this affidavit.

7 INTRODUCTION

8 I have been asked by you to opine about the so-called "Flynn Effect" and
9 its use in correcting IQ scores in so-called "Atkins" criminal proceedings intended
10 to consider whether a defendant has mental retardation (MR). The two
11 purposes of my involvement, as I understand it, are to educate the court about
12 this phenomenon (which is confusing to most non-psychologists and even to
13 many psychologists) and to consider whether the Flynn correction formula was
14 applied correctly by your expert, Dr. Ginger Calloway, in the case of your client
15 Antonio Chance. Although I have not been asked, I am also available to opine
16 about the nature of MR and the criteria used to diagnose it. This is a role that I
17 have played in various other cases, and it is a topic on which I am considered a
18 leading authority.

19 In the first section below, I shall summarize my qualifications in the fields of
20 mental retardation and psychology, and note my expertise on the Flynn Effect.
21 In the second section, I shall briefly describe the Flynn Effect and indicate why it
22 is important to take to into account when diagnosing MR. In the third and final
23 section, I shall opine about whether the Flynn correction formula was applied
24 correctly in this case.

25 MY BACKGROUND

26 In the past four years, I have been qualified as an expert on MR and related
27 cognitive disorders in four or five capital proceedings in the states of Arizona, California

28 and Colorado. In addition, I have previously been qualified as an expert on MR in family
29 court proceedings in New Jersey and Connecticut. I am a licensed psychologist in the
30 state of Nebraska and was previously licensed in the state of Tennessee (current status:
31 inactive). In addition to testifying in several so-called “Atkins” proceedings, I have been
32 a consultant (and submitted declarations) in numerous other cases. Although my work,
33 with two exceptions, has generally been at the request of attorneys representing
34 defendants, I have found that a claim of mental retardation was unjustified in
35 approximately half of the cases in which I actually examined a defendant (in contrast to
36 other cases, in which my role was limited to educating the court about the nature of
37 mental retardation and/ or opined about the adequacy of reports by other experts.)

38 I am a Clinical Professor of Psychiatry at the University of Colorado Health
39 Sciences Center, and Emeritus (retired) Professor of Educational Psychology at the
40 University of Connecticut. I received a Ph.D. in Developmental Psychology from the
41 University of Rochester, and was a Postdoctoral Fellow in Mental Retardation and
42 Developmental Disabilities at UCLA’s Neuro-psychiatric Institute. Before moving to
43 Connecticut, I held academic appointments at the University of Nebraska and at George
44 Peabody College of Vanderbilt University.

45 I have been elected “Fellow” (a designation given only to the most qualified
46 members) by the Mental Retardation division of the American Psychological Association
47 and by the American Association on Mental Retardation (now the American Association
48 in Intellectual and Developmental Disabilities). I was also elected to a term as President
49 of the Academy on Mental Retardation, which is the most prestigious research
50 organization in the field. I have published extensively on MR, with particular emphasis
51 on “adaptive behavior.” I am a leading scholar in the MR field, as seen in the most recent
52 diagnostic manual of the American Association on Mental Retardation [MENTAL
53 RETARDATION: DEFINITION, CLASSIFICATION AND SYSTEMS OF SUPPORTS
54 (10th Edition, 2002), hereinafter “the 2002 AAMR Manual”], which cited at least twelve
55 publications by me, more than that of any other authority. My book WHAT IS MENTAL
56 RETARDATION, co-edited with H. Switzky (AAMR; 2003; rev. ed. 2006) has, in a

57 short time, become one of the most-quoted reference works in the field of mental
58 retardation and has been described by Yale professor Edward Zigler as “the best book
59 ever written about the definition and diagnosis of mental retardation.” In 2008, AAMR/
60 AAIDD recognized my lifetime of contributions to the field by granting me its highest
61 honor, the Gunnar and Rosemary Dybwad Award for Humanitarianism.

62 In the course of my testimony in criminal cases, I have on occasion been asked to
63 explain and comment on the so-called “Flynn Effect”. It is a topic that I have written
64 about in two published papers: Stephen Greenspan (2006), Issues in the use of the
65 “Flynn effect” to adjust IQ scores when diagnosing MR. Psychology in
66 Mental Retardation and Devel.Disabilities, 31 (3), 3-7; and Stephen
67 Greenspan, S. (2007), Flynn-Adjustment is a Matter of Basic Fairness.
68 Psychology in Mental Retardation and Developmental Disabilities. In
69 addition, I was the editor of a special section on Flynn in 2007 in the
70 above-named journal, and edited a paper by Dr. Flynn that appeared
71 alongside mine. Currently, I am the editor of a forthcoming special issue
72 devoted to Atkins matters of the journal Applied Neuropsychology and Dr.
73 Flynn has agreed to participate in that as well. I am in frequent contact
74 with Dr. Flynn and it is fair to say that among psychologists specializing in
75 mental retardation, few know more about the Flynn Effect than I do.

76 A BRIEF EXPLANATION OF THE “FLYNN EFFECT”

77 James Flynn is an American citizen who lives in New Zealand and is
78 an Emeritus Professor at New Zealand’s prestigious University of Otago. All
79 of the data that he has used as the basis for the Flynn Effect calculation
80 involves Americans who were tested in the US on tests published in the US.
81 I say this because I was once asked by a judge “why should we care
82 about what a New Zealander studying New Zealanders has to say about
83 American intelligence?” Furthermore, the Flynn Effect has been replicated
84 by many US researchers and has been accepted as valid by the

85 American Psychological Association (which devoted a monograph to the
86 Flynn Effect) and by the International Society for Intelligence Research,
87 which presented its highest award to Dr. Flynn at its annual meeting in
88 Albuquerque in 2005.

89 To understand the Flynn Effect, one has to understand that all IQ
90 tests are “norm-referenced”, which means that a score places a subject
91 at a specific percentile of the norming population. In the case of major IQ
92 tests such as the WAIS, the norming population is intended to reflect the
93 population of the United States. This is a painstaking process, in which the
94 2,000-plus members of the norming sample are very carefully assembled
95 (e.g., with comparable percentages according to gender, ethnicity,
96 education level, region, etc) to be fully representative of the US
97 population as a whole. In the case of the WAIS-III, the reference
98 population framework used to compile the sample was the 1995 US
99 census. It is because compiling a representative standardization sample is
100 such a cumbersome and expensive process, that IQ tests are so pricey
101 and so few tests of intelligence are considered acceptable when
102 diagnosing MR (which attests that an individual's intelligence is at or
103 below approximately the second percentile of the population in terms of
104 intelligence).

105 Because every edition of the WAIS and other Wechsler scales are
106 carefully standardized in relation to its current US population, one is on
107 safe ground in assuming that the percentiles obtained from each edition
108 are valid reflections of the U.S. population at the time of standardization.
109 For each WAIS and other IQ test, a raw score at exactly the population
110 mean is assigned an IQ score of 100 (which means a subject is at the 50th
111 percentile). Similarly a raw score at the second percentile is equivalent to
112 a standard score of 70. This is because all IQ scores use a standard

113 deviation of 15, and the second percentile of the population is
114 established at minus two standard deviations [$100 - (15 \times 2) = 70$].

115 In order to ensure that every edition of the WAIS is comparable to
116 every other edition, the raw score mean at the time of test construction is
117 established at 100. James Flynn's great contribution (which is considered
118 the most important advance in the past several decades in intelligence
119 research) was to discover that over the past two generations, the norms
120 are progressively toughened, because of better average performance by
121 the standardization sample. Flynn discovered this by looking at a large
122 number of studies, some done by the test publishers themselves, in which
123 a sample of subjects take both the new edition, and its predecessor
124 edition, of the same IQ test. An example is the WISC-R, published in 1974,
125 and the WISC-3, published 17 years later in 1991. Subjects taking the two
126 tests (at the same time or a year or two apart) scored on average over
127 five points lower on the WISC-III than they did on the predecessor test, the
128 WISC-R. The only explanation for this, given that the individuals did not
129 change in such a short time span, is that the test norms were made more
130 difficult. In other words a raw score that would have placed a subject at
131 the population mean on the old test would put the same individual below
132 the population mean on the new test. The only plausible explanation for
133 that is that the average person in the population at the time of the
134 standardization of the new test is smarter (at least as measured by an IQ
135 test) than the average person in the population at the time of the
136 standardization of the previous edition of that test.

137 IQ tests are revised and reissued every fifteen years or so. The
138 above-described phenomenon, whereby norms are toughened for each
139 succeeding generation of tests to reflect improved average performance
140 of the standardization sample, has been found to apply to all of the

141 major tests—such as the versions of the WAIS, WISC, and Stanford-Binet--
142 for both children and adults. The proof of this is in the kind of study cited
143 above, where subjects taking the old and new version of a test show a
144 drop in IQ score on the new test which, when divided by years of norm
145 obsolescence, averages about 3.3 points per decade (or 0.33 points per
146 year). Again, the only explanation for this drop of performance is that
147 norms have been toughened to reflect a higher level of performance by
148 the new standardization cohort. Given that the major test publishers--such
149 as The Psychological Corporation (now part of Harcourt Assessment/
150 Pearson), publisher of the WAIS and WISC—go to great trouble to ensure
151 that the standardization sample is representative of the US census at the
152 time of test development, the only explanation for the improved
153 performance of the standardization sample is that the US population has
154 gotten “smarter” in the elapsed time between development of the old
155 and new editions.

156 There has been no credible doubting of the reality of the Flynn
157 Effect within the psychology field, although the reasons for this
158 phenomenon are still open to debate. (An examination of the various
159 hypotheses that have been put forth is beyond the scope of this
160 discussion). Because the Flynn Effect refers to changes in the intelligence
161 of the standardization cohort (and the broader population which the
162 standardization sample represents), the Flynn Effect applies to any use of
163 IQ testing, including diagnosing mental retardation. This is illustrated in a
164 recent study [T. Kanaya, M.H. Scullin & S.J. Ceci (2003). The Flynn effect
165 and US policies: The impact of rising IQ scores on American society via
166 mental retardation diagnoses. *American Psychologist*, 58, 778-790] in
167 which the investigators looked at changes in IQ and mental retardation
168 diagnoses in children, according to which version of an IQ test was used.

169 Because children receiving special education must be retested every
170 three years, Kanaya et al looked at children who were tested and
171 retested during the period when schools were switching over from the
172 older WISC-R to the newer WISC-3 IQ test. They found that the same
173 children who were previously found not to have mental retardation on the
174 WISC-R were much more likely to be given a diagnosis of mental
175 retardation when tested on the new WISC-3, because their IQ scores
176 dropped an average of five points on the WISC-3, a decline attributable
177 to the same 0.33 points rise per year in the performance of the
178 standardization sample (over the 17 years between the two test editions)
179 noted earlier.

180 The case for making "Flynn adjustments" in IQ scores when
181 diagnosing mental retardation for any purpose, but especially in
182 something as serious as a capital proceeding, is that it should not be a
183 lottery whose outcome hinges on which test a psychologist uses and
184 when he or she uses it. Flynn adjustment is, thus, an attempt to create a
185 level playing field, in which all petitioners are held to the same standard,
186 by correcting for the confound of test obsolescence. The convention that
187 has been adopted in numerous court cases with which I am familiar is that
188 an obtained IQ score is adjusted downward by determining the age of
189 the test edition that was used and then subtracting 0.33 points for every
190 year of norm obsolescence. Thus if a subject receives a full-scale IQ score
191 of 72 on a test with norms that are 10 years old, one multiplies 10 by 0.33,
192 with the resulting product of 3 (3.3 rounded down) subtracted from 72,
193 producing a Flynn-adjusted score of 69.

194 WAS THE FLYNN EFFECT CALCULATED CORRECTLY BY DR.
195 CALLOWAY?

196 I was asked to express a judgment regarding whether Defense
197 psychologist Dr. Ginger Calloway followed the proper procedure in
198 adjusting IQ scores obtained from Mr. Chance. I was provided a Table
199 (see below) that shows all of what are labeled “statute approved” IQ tests
200 for Mr. Chance, the name of each test, the date of administration, the
201 date when each test was normed, the source of the testing, the full-scale
202 IQ that was obtained, and the Flynn-adjusted IQ she obtained. This is the
203 only data provided me from Dr. Calloway’s evaluation. I have not seen
204 her report, or anything else about Mr. Chance or the case.

205 The WPPSI is the Wechsler Preschool and Primary Scale of
206 Intelligence, and is intended for use before the age of six. The WISC-R is
207 the Wechsler Intelligence Scale for Children—Revised, and is intended for
208 use between the ages six and sixteen. The WAIS-III is the Wechsler Adult
209 Intelligence Scale—third edition, and is intended for use between the
210 ages sixteen and, I believe, eighty-nine. The WASI is the Wechsler
211 Abbreviated Intelligence test, a short-form version of the WAIS-III that is
212 intended for use with the same age range. Unlike the WAIS-III which is
213 administered with ten or more sub-scales, the WASI contains only four of
214 the sub-scales and a full-scale IQ can be obtained from either two or all
215 four of these sub-scales. It is not clear from the information given to me
216 which configuration of the WASI was used by the state-appointed
217 psychologist who administered the WASI, although I assume it was the four
218 sub-scale configuration.

219

220

221

222

223 Table showing Dr. Calloway's Flynn Adjustments on A. Chance's IQ scores

	A	B	C	D	E	F	G
1							
2	<u>Adjusted Scores (Flynn adjusted for test obsolescence)</u>						
5	<u>STATUTE ACCEPTED TESTS</u>						
6	<u>TEST</u>	<u>Admin Date</u>	<u>Obtained Score</u>	<u>Examiner</u>	<u>Adjusted IQ</u>	<u>Test Standardization Date</u>	<u>Test Age</u>
7	WPPSI	1983	69	School I	63.67	1967	16
8	WISC-R	1987	75	School I	70.00	1972	15
9	WISC-R	1990	74	School I	68.00	1972	18
10	WAIS-III	2006	67	Defense	63.67	1996	10
11	WASI	2007	66	State	63.00	1998	9

224

225

226 I was not provided a definition of “statute approved” tests, but I assume that this
227 term refers to tests that have adequate reliability and validity, including valid norms, for
228 diagnosing mental retardation for legal or government administrative purposes. There are

229 only a few tests (which includes the Stanford-Binet-5) that are given this designation, and
230 the various Wechsler Scales are considered the “gold standard” of IQ tests. As a rule,
231 short-form IQ tests are not statute-approved, for the simple reason that the larger the
232 number of sub-scales the more accurate the full-scale IQ score is considered to be. This
233 can be seen in a research article [Bradley N. Axelrod (2002). “Validity of the Wechsler
234 Abbreviated Scale of Intelligence and other very short forms of estimating intellectual
235 functioning”, Assessment, vol. 9, No.1, 17-23] in which questions were raised about the
236 accuracy of the scores obtained from the WASI, when compared to the full WAIS-III.
237 These concerns had more to do with the performance and verbal IQ scores (obtained from
238 only one or two sub-scales on the WASI) than from full-scale IQ estimates, but even so
239 this shows why one should generally not rule MR in or out from a short-form IQ test such
240 as the WASI.

241 Generally, short-form IQ tests are used as screening instruments, to see if more
242 extensive testing is needed, or to confirm the results obtained from other testers. This last
243 reason likely explains why the state’s psychologist administered the WASI to Mr.
244 Chance, namely to see if the MR-qualifying IQ score of 67 obtained a year earlier by the
245 defense psychologist was believable. As the WASI produced a nearly identical score of
246 66, he or she likely concluded that more extensive IQ- testing was not needed.

247 Another question I have about the WASI has to do with the standardization date
248 of 1998 that is listed in the table, as contrasted with the standardization date of 1996 for
249 the WAIS-III. As the WASI used WAIS-III sub-scales, and its norms were presumably
250 obtained from the same subjects, the fact that the WASI may have been published two
251 years subsequent to the WAIS-III does not justify using a later standardization date. For
252 this reason, a Flynn adjustment of the WASI should probably use the 1996
253 standardization date, which would actually bring the score down by an additional 0.66
254 point. However, I am not certain that it is appropriate to use the 0.33 adjustment figure
255 for the WASI, as Flynn studies of short-form tests have generally not been done. For this
256 reason, it is probably more appropriate to use the obtained 66 rather than the adjusted 63
257 for the WASI.

258 If one takes into account the five point confidence interval for the Wechsler
259 scales, as recommended in the MR classification sections in the diagnostic and
260 classification manuals for both DSM and AAMR, then every one of the five IQ tests
261 reported in the Table qualify Mr. Chance for the first (IQ) prong of the official North
262 Carolina definition of MR. That is because the highest score is 75, which is at the upper
263 boundary of the category taking into account the 70-75 standard error-adjusted range.
264 However, if one adheres (not recommended, in my opinion) to a rigid upper boundary of
265 70, then the two WISC-R's (in 1987 and 1990) are slightly above the boundary. Flynn
266 adjustments (justified, as the WISC-R was over 15 years obsolescent) bring those scores
267 to at or just below the 70 ceiling.

268 An example of how one would do a Flynn adjustment follows, using the 1990
269 WISC-R as illustration. The score obtained in 1990 on Mr. Chance was 74. The WISC-R,
270 standardized in 1972, was 18 years obsolete in 1990. Multiplying 18 years of norm
271 obsolescence by the annual change rate of 0.33, one obtains a sum of 5.94, which is
272 rounded to six. Subtract this amount from the obtained score of 74, and one gets a Flynn-
273 adjusted IQ score of 68, which is the number shown in column E in the Table.

274 I have checked all of the Flynn-corrected scores in the table provided from Dr.
275 Calloway, and all of the scores check out as correct. As mentioned above, I did not do
276 this for the WASI, as I do not consider a short form test to be appropriate for diagnosing
277 MR. I also did not do this for the WPPSI, as I do not believe the WPPSI or any other pre-
278 school or infancy IQ measure was studied by Dr. Flynn, and I have never before seen a
279 Flynn adjustment done for the WPPSI. As a rule, MR is not a diagnosis made for pre-
280 schoolers (except for the most severely and multiply-handicapped cases, where IQ tests
281 are not really needed). That is both because preschool IQ tests are not as reliable as child
282 or adult IQ tests, but also because preschoolers often experience cognitive growth spurts
283 and one is reluctant to assign so serious a label as MR to very young children. So I
284 preferred not to do a Flynn adjustment for the 1983-obtained WPPSI. However, as
285 indicated, the obtained score of 69 is under the MR ceiling, even without a Flynn
286 adjustment.

287 In sum, all of the five IQ scores for Mr. Chance are in the MR ballpark of a
288 ceiling score of 70-75, even without a Flynn adjustment. The two WISC-R scores that are
289 above the absolute ceiling of 70 are brought into compliance with that ceiling if one (very
290 justifiably) uses the Flynn adjustment factor of 0.33 for the many years of obsolescence
291 of the WISC-R at the time of testing. In my view, Dr. Calloway conformed with
292 appropriate and accepted practice for making these adjustments, as can be seen in the
293 above illustration.

294

295 Everything I have sworn to in this affidavit is my own words and is honest and
296 truthful.

297

298 FURTHER AFFIANT SAYETH NOT

299

300

301

Stephen Greenspan

302

303

Sworn to or affirmed before me this 25th day of April, 2008

304

305

306

Notary Public

307

308

My commission expires:

309

310

311

