Simple reaction time: It is not what it used to be

IRWIN W. SILVERMAN
Bowling Green State University

This article calls attention to the large amount of evidence indicating that simple visual reaction time (RT) has increased. To show that RT has increased, the RTs obtained by young adults in 14 studies published from 1941 on were compared with the RTs obtained by young adults in a study conducted by Galton in the late 1800s. With one exception, the newer studies obtained RTs longer than those obtained by Galton. The possibility that these differences in results are due to faulty timing instruments is considered but deemed unlikely. Of several possible causes for longer RTs, two are regarded as tenable: that RT has been increased by the buildup of neurotoxins in the environment and by the increasing numbers of people in less than robust health who have survived into adulthood. The importance of standardizing tests of RT in order to enable more refined analyses of secular trends in RT is emphasized.

The study of reaction time (RT) has a long history, dating back to the mid-19th century (Boring, 1950). Over this period of time, a number of human attributes have markedly changed in either prevalence or magnitude. Examples include life expectancy (Riley, 2001), height (Cole, 2000), weight and obesity (World Health Organization, 2008), myopia (Rose, Smith, Morgan, & Mitchell, 2001; Storfer, 1999), running speed (Tomkinson, Léger, Olds, & Cazorla, 2003), and IQ (Flynn, 1984, 1987). The primary purpose of this article is to call attention to the large amount of evidence indicating that simple visual RT has also changed markedly over time. A secondary purpose is to evaluate a number of possible explanations for the change in RT documented here.

To show that simple visual RT has changed over time, the RTs obtained in a study conducted by Francis Galton in the late 1800s will be compared with the RTs obtained in subsequent studies. The decision to use Galton's study to provide reference data for making historical comparisons was based on the fact that relative to the samples tested by his contemporaries, the sample tested by Galton was many times larger and far more diverse in terms of socioeconomic status. Additionally, Galton tested RT in both men and women, whereas in the RT studies conducted by his

American Journal of Psychology
Spring 2010, Vol. 123, No. 1 pp. 39–50 •

contemporaries, the samples were often composed of men only.

Galton's study of RT

The story of how Galton came to study RT was told by Boring (1950). Briefly, Galton's interests in individual differences and measurement led him to establish what he called the Anthropological Laboratory at two sites in London, England, first at the International Health Exposition and then at the South Kensington Museum. Over the life of the Anthropological Laboratory, from 1884 to 1893, more than 17,000 people, for a small fee, were measured on an assortment of physical, sensory, and psychomotor attributes. Among the attributes measured was simple RT to light and sound, which Galton supposed might reflect general mental ability (Jensen, 1980). (Additional details about the test battery used by Galton can be found in Johnson et al., 1985.)

There are both knowns and unknowns about the method Galton used to measure visual RT. We know that RT was measured with a pendulum chronoscope designed by Galton (1889) himself; that when the pendulum was released, this caused a mirror to deflect, thereby throwing a light on or off a screen (Koga & Morant, 1923), and that participants pressed a key in response to the stimulus (Galton, 1889). However, we do not know whether participants responded to the onset or offset of the light, although it was probably the former because RT to sound was measured by having a weight hit a hollow box. We also do not know whether participants completed practice trials and how many test trials were given. However, regarding the latter, Johnson et al. (1985) stated, "As far we can ascertain, Galton took only a single measure each of visual and auditory reaction times" (p. 879). Still another unknown is whether there was a warning signal, but it would seem hardly necessary to give a warning signal if there was only one test trial for each stimulus.

Galton died in 1911, never having published a description of the procedure used to test RT or an analysis of the RT data that were collected. However, thanks to Koga and Morant (1923), we have the mean RTs for more than 3,000 males tested at both sites, and thanks to Johnson et al. (1985), we have the mean RTs for more than 4,000 males and 1,000 females tested at the South Kensington Museum. Only the data reported by Johnson et al. are used here, as they include both sexes. Also, even though the RTs obtained by Galton cover a wide age range, only the data for young adults are considered here, because the literature search for newer studies yielded far more data for young adults than for any other age group.

It appears that Galton tested RT in each modality on a single trial. If so, there is reason to doubt whether RT was measured reliably. Johnson et al. (1985) cited an unpublished paper (Macey, Plomin, & McClearn, 1979) that reported the test–retest reliabilities for a sample of 50 people who were retested within 1 month on the battery of tests assembled by Galton. Johnson et al. presented no exact reliabilities, but they stated, "In general, the reliabilities were quite high" (p. 878). Beyond these unpublished results, there are two reasons for believing that RT was measured reliably in Galton's study. One is that the visual and auditory RTs obtained by Galton were significantly correlated. As computed by Johnson et al., the correlations were .50 and .54 for men and women, respectively, at ages 26 and older. In a later study by Forbes (1945), the RT to light and to sound was tested on a large (though varying) number of trials in men ages 17 to 53. The correlation between the RTs for the two modalities was .43. Interestingly, despite Forbes having tested RT on far more trials than did Galton, the preceding correlation is significantly smaller than the correlation for men in Galton's study, $z = 13.93$, $p < .001$. Another reason for believing that RT was measured reliably in Galton's study is that RT was correlated in siblings. As computed by Johnson et al., the correlations between the visual RTs for several sibling pairings ranged from .30 (for first son and second son) to .45 (for first son and first daughter). These sibling correlations compare favorably with the sibling correlations on several physical attributes. For example, the correlations between the weights of siblings for the two sibling pairings just mentioned were .35 and .19, respectively. (The sibling correlations in all cases were corrected for age.)

Table 1 presents the mean RTs for Galton's sample of men and women from ages 18 to 30 as reported by Johnson et al. (1985). No standard deviations are given in this table because none were provided by Johnson et al. (Koga & Morant reported a standard deviation of 40.03 ms for men, but it was for the entire age range tested by Galton, 5.5 to 81.5 years). The total *N*s were 2,522 and 302 for the men and women, respectively. Given that the RTs varied over a narrow range for the two genders (8 ms for men and 9 ms for women), the median RTs, 183.0 ms for men and 187.0 for women, were used in the comparisons to be described. (Alternatively, for comparative purposes, the weighted means for each sex could have been used, but these were identical to the median RTs.)

The comparison studies

This article builds on a meta-analysis by Silverman (2006) in which sex differences in simple visual RT were examined. However, whereas Silverman included in the meta-analysis only studies in which both sexes were represented, the present review included studies in which one or both sexes were represented. To locate studies for this article, a search was conducted of the PsycINFO and Medline databases using as key words *reaction time, simple reaction time, neuropsychological assessment,* and *neurobehavioral assessment*. Also searched were two reviews of the RT literature (Teichner, 1954; Woodworth & Schlosberg, 1954) and the reference lists of the retrieved studies. The following criteria were used in selecting studies for inclusion in the present review. First, the sample consisted of people recruited from the general population and whose ages ranged from about 18 to 30 years. Second, the study sample had not been exposed to high levels of neurotoxins, substances known to increase RT. Third, given that Galton's sample was British, the study was conducted in a Western country. Fourth, the study sample had to be 20 or larger in size for each gender. Fifth, the delivery of the stimulus was not predictable, which ruled out studies in which the interval between stimuli was fixed or increased or decreased according to a regular pattern. Sixth, the response to the stimulus was manual in nature, such as pressing or releasing a button or key. Seventh, to effect the response, the arm did not have to be moved. (This restriction was based on the consideration that if the arm must be moved, RT is necessarily lengthened.) Eighth, the RT measure was representative of the total set of RTs. This restriction eliminated studies in which RT was measured in terms of the best RT or the longest or shortest RT.

No restriction was placed on studies with respect to whether a warning signal preceded the stimulus. Teichner (1954) stated that Wundt found that a warning signal decreased RT, but no independent corroboration could be found for this result, at least not when the stimulus followed the warning signal at varying and unpredictable intervals. Indeed, contrary to Teichner, data reported by Seashore, Starmann, Kendall, and Helmick (1941) show that RT did not differ as a function of whether a warning signal preceded the stimulus, $t(46) = .45$, *ns*.

Table 2 lists the studies in which the data were presented separately for men and women or in which only one sex was represented. The table reports for each study the country in which the study was conducted, the age and number of participants, the mean and standard deviation for the RT measure, and the 95% confidence interval for the mean RT. The notes to the table report the source of participants, number of practice and test trials, and special treatment (if any) of the data. It should be noted that Lefcourt and Siegel (1970) did not report standard deviations for their data. The pooled standard deviation for this study was estimated from an analysis of variance table in the study report using a formula given by Lipsey and Wilson (2001).

Two other studies were retrieved in which RT data were presented for the two genders combined. One study was conducted by Wilkinson and Allison (1989), who measured simple visual RT in a sample

TABLE 1. Descriptive statistics for simple visual reaction times obtained in Galton's study

	Men		Women	
Age	N	M	N	M
18	363	181	73	187
19	368	182	98	187
20	345	183	85	184
21	268	185	81	190
22	215	180	81	185
23	183	188	80	181
24	141	184	78	185
25	126	188	67	188
26–30	513	181	235	189

Source. Johnson et al. (1985).

TABLE 2. Descriptive statistics for simple visual reaction times obtained in the comparison studies

			Men			Women			
Study	Country	Age	*N*	*M* (*SD*)	95% CI	*N*	*M* (*SD*)	95% CI	Notes
Anger et al. (1993)	United States	16–25	80	260 (30)	253–267	140	285 (40)	278–292	Sample composed of postal, hospital, and insurance workers living in three cities. No. of practice trials = NR; no. of test trials = 64. RT estimated from a graph.
		26–35	73	250 (25)	244–256	163	280 (30)	275–285	
Brice & Smith (2002)	Great Britain	19–23	24	324 (43)	307–342	—	—	—	Sample composed of students. No. of practice trials = NR; no. of test trials ~ 60.
Deary & Der (2005a)	Great Britain	24	254–255	295 (78)	285–305	288–289	306 (73)	298–314	Sample recruited from a random sample of people living in Scotland. Test given in 1987–1988. No. of practice trials = 8; no. of test trials = 20.
Der & Deary (2006)	Great Britain	18–30	834	300 (84)	291–309	1,023	318 (96)	312–324	Sample recruited from random sample of British adults. Test given in 1984–1985. No. of practice trials = 8; no. of test trials = 20.
Forbes (1945)	Great Britain	17–36	76	286 (39)	278–295	—	—	—	Sample described as "fairly representative of male community." No. of practice trials = NR; no. of test trials = NR.
Jorm et al. (2004)	Australia	20–24	1,163	214 (29)	212–216	1,241	224 (36)	222–226	Sample recruited from a random sample of registered voters in Canberra and neighboring town. Test given in 1999–2000. No. of practice trials = NR; no. of test trials = 80. RTs >2,000 ms and RTs ≥3 *SD*s from each participant's *M* for each block of 20 trials excluded.
Lefcourt & Siegel (1970, Exp. 1)	Canada	22	40	236 (15)	231–241	40	263 (15)	258–268	Sample composed of students. Data given here are for two conditions jointly considered; irregularly varied preparatory intervals between warning signal and stimulus and trial onset experimenter controlled. No. of practice trials = NR; no. of test trials = 60.
Reed et al. (2004)	Canada	18–25	171	253 (30)	249–257	198	268 (39)	263–273	Sample composed of students. No. of practice trials = NR; no. of test trials = NR.

TABLE 2. (cont.)

Study	Country	Age	Men N	Men M (SD)	Men 95% CI	Women N	Women M (SD)	Women 95% CI	Notes
Seashore et al. (1941)	United States	18–22	47	197 (23)	190–204	—	—	—	Sample composed of students. No. of practice trials = 25. Warning signal given. No. of test trials = 100.
				203 (23)	196–210	—	—	—	Same sample as above. Procedure as above except no warning signal given.
Smith et al. (1999)	Great Britain	18–22	64	306 (44)	295–317	—	—	—	Sample drawn from a volunteer pool. Data here are for participants who did not contract an upper respiratory tract illness over course of study. No. of practice trials = NR; no. of test trials ~ 60.
Taimela (1991)	Finland	20	893	199 (46)	196–202	—	—	—	Sample composed of military recruits. Tested in 1990. No. of practice trials = NR; no. of test trials = 5.
Taimela et al. (1991)	Finland	20	123	183 (23)	179–187	—	—	—	Sample composed of military recruits. Tested in 1987. No. of practice trials = NR; no. of test trials = 5. Each participant's longest and shortest RTs excluded.

Note. *CI = confidence interval; NR = not reported.*

of visitors to a science museum located in London, England. There were 10 trials, but the first 2 were treated as practice. RTs were not retained for some unspecified number of participants because they had five or more RTs greater than 1,000 ms or because they had one or more RTs greater than 2,500 ms. For each of the remaining participants, any RT less than 120 ms was considered to be an outlier and was deleted from the participant's RT distribution. A total of 1,l89 people were tested at ages 20–29. In the other study (Krieg et al., 2001), which was conducted from 1988 to 1994, the participants were a nationally representative sample of adults living in the United States. The test of simple visual RT consisted of 50 trials, but the first 10 trials were treated as practice. Also not included in the data analysis were RTs less than 50 ms or greater than 750 ms. A total of 1,526 people were tested at ages 20–29.

Results

Table 2 presents the means and standard deviations for the comparison studies in which participants were of one sex or in which participants were of both sexes and the results were presented separately for each sex. The total *N* is 3,836 for men and 3,093 for women. Overall, the mean RT was 250.43 ms (*SD* = 46.53) for men and 277.71 ms (*SD* = 30.76) for women. However, as seen in the table, in both sexes the mean RT varied widely across studies, ranging from 183 to 324 ms for men and 224 to 318 ms for women.

Recall that in Galton's study the median RTs obtained by men and women between ages 18 and 30 were 183.0 and 187.0 ms, respectively. Consider now the 95% confidence intervals for the mean RTs given for the comparison studies in Table 2. Note that for men there is only one study (Taimela et al., 1991) in which this interval included 183.0 ms, and for women there is no study in which this interval included 187.0 ms. Moreover, with the exception noted, the RTs obtained in the comparison studies were all longer than the RTs obtained by Galton.

Consider next the two studies in which the data were not reported separately by sex. In the study by Wilkinson and Allison (1989), the mean RT was about 245 ms (as estimated from a graph) at ages 20–29. No standard deviation is reported by Wilkinson and Allison, making it impossible to compute a 95% confidence interval for their mean RT. But undoubtedly it would not have included the median RT for men or women in Galton's study. In the study by Krieg et al. (2001), the mean RT was 230.70 ms ($SD = 44.75$). The 95% confidence interval for the mean ranged from 228 to 233 ms, which does not include the median RT for either men or women at ages 18–30 in Galton's study. Thus, both Wilkinson and Allison's and Krieg et al.'s data also indicate that RT has increased over time.

As noted earlier, Galton's sample was diverse in terms of socioeconomic status. Johnson et al. (1985), divided the men tested into seven occupational groups and the women tested into six occupational groups. Then they subdivided each of these groups into two age levels, 14–26 and 26+. For men the RTs for each of these occupational and age groups ranged from 173 to 195 ms, and for women they ranged from 164 to 213 ms. Because the longest RT for women was based on only three participants, it cannot be regarded as representative, but because the next longest RT, 207 ms, was based on 41 participants, it is probably representative. Inspection of Table 2 shows that there is only one comparison study (Taimela et al., 1991) in which the mean RT for men fell within the range of the RTs obtained by the men in Galton's study and no comparison study in which the mean RT for women fell within the range of RTs obtained by the women in Galton's study. Thus, the differences in results obtained by Galton and those obtained in the comparison studies cannot be readily explained by differences in the socioeconomic backgrounds of the participants.

A final question addressed is whether for the comparison studies the RTs for any specific socioeconomic group deviated from the general finding that RT has increased over the past 100 or so years. Only one of the comparison studies reported results separately for different income or education levels. In this study by Krieg et al. (2001), RT in both sexes combined decreased as education and income increased. However, the shortest RT by education was 225.23 ms (at grade 13+ of education) and the shortest RT by income was 224.31 ms (at ≥$50,000 of income), and both of these RTs are significantly ($ps < .001$) longer than the median RTs for Galton's sample of men and women at ages 18–30. Unfortunately, Krieg et al. did not break the data down further so that it can be determined how RT varied as a function of education and income within each sex at different age levels. Nonetheless, it appears that the conclusion that RT has increased over the past 100 or so years does not have to be qualified with regard to socioeconomic status, although the rate of change could have differed as a function of socioeconomic status.

Discussion

The present review provides strong evidence that RT has increased in both sexes since Galton conducted his landmark study in the late 1800s. Furthermore, the present review indicates that the increase in RT has occurred across the entire socioeconomic spectrum and in at least four out of five Western countries.

Two questions cannot be answered with the present dataset: When did the increase in RT begin, and at what rate did it increase over time? As to the first question, all that can be said is that RT began increasing before the 1940s, but how much earlier than this decade cannot be said because the present dataset contains no studies that were published between 1894 and 1941. As to the second question, a definite answer cannot be given because the present studies differ with respect to several variables that could affect RT: composition of the samples, testing procedure, and handling of data. In principle, it would be possible to uncover the rate at which RT increased by controlling for potentially confounding variables in a multiple regression analysis. However, this requires that each of these variables be represented by multiple data points, but this requirement cannot be met by the present dataset. Accurately describing change over time also requires that both ends of the temporal dimension be well represented in the dataset and that the dataset be free of outliers (Cohen, Cohen, West, & Aiken, 2003); neither of these requirements can be met in the present dataset. Thus, it is important to reiterate that the purpose of this review is not to show that RT has changed according to a specific function over time but rather to show that modern studies have obtained RTs that are far longer than those obtained by Galton.

The present data indicate that although RT has increased in both sexes, it has increased more in women than men. Whereas in Galton's study the RTs for women were on average only 4 ms longer than those for men, in the six comparison studies in which both men and women were tested, the RTs for women were on average more than 17 ms longer than those for men. This widening of the sex difference over time should be viewed with caution, however, because the women in Galton's study may have been somewhat atypical. Suggesting this possibility is the fact that there were 2.9 times as many men as women in Galton's sample of young adults. This could have come about if many more men than women visited the South Kensington Museum. This is highly doubtful because at ages 11–17 there were 3.8 times as many boys as girls in Galton's sample, and it seems improbable that parents were far more likely to bring their sons than daughters to the museum. A more reasonable explanation for the sharp imbalance in the numbers of males and females in Galton's sample is that at that time females were more reluctant to be tested than were males. (After all, Galton conducted his study during the Victorian era, when the feminine ideal was to be weak, helpless, and demure). If this is correct, then it could be that in Galton's sample of females there was an overrepresentation of females with above-average physical and sensory abilities. As will be seen, the males in Galton's sample obtained RTs similar to those obtained by the participants in other early RT studies, most of whom were male. This suggests that, in contrast to the females, the males were typical of males at that time.

The issue of instrumentation

The conclusion that RT has changed over time would be rendered moot if the apparatus Galton used to measure RT was defective. More specifically, it might be argued that there was a lag between the release of the pendulum and the presentation of the stimulus, thus resulting in RTs that were too short by some constant duration. Perhaps the strongest counterargument to this objection is that the RTs reported by other early RT researchers are quite comparable to those obtained in Galton's study. This is seen in an early review of the RT literature by Ladd and Woodworth (1911) covering eight studies (not including Galton's). These RTs, which are given by Robinson (1934), ranged from 151 ms to 200 ms, with the median RT being 192 ms. Note that the RTs obtained by Galton fall within the range of RTs obtained in these early RT studies. None of these early researchers measured RT using an apparatus identical to the one used by Galton, and in fact some of them used the Hipp chronoscope to measure RT, a timing instrument based not on the pendulum but on a clockwork mechanism (Woodworth & Schlosberg, 1954). So it is highly unlikely that the RTs obtained by Galton are simply artifacts of the apparatus used to measure RT.

Why has simple visual RT increased?

There are at least four explanations for the increase in RT documented here. The first explanation is that as people have grown taller over the past 150 years (Cole, 2000), RT has increased because the nerve impulse has to travel farther as the body lengthens. However, contrary to this explanation, Galton's data (see Johnson et al., 1985) show that taller adults had slightly shorter RTs than did smaller adults. His data also show that adults with longer arm spans and longer left arms also tended to have slightly shorter RTs than did adults with shorter arm spans and shorter left arms. Additional evidence indicating that the increase in RT cannot be accounted for by increases in body size comes from two contemporary studies. One study, with Australian adults, found height to be unrelated to RT when age, gender, and education were controlled (Anstey, Dear, Christensen, & Jorm, 2005), and the other study, with American adults, found no relationship between RT and height (Kilburn, Thornton, & Hanscom, 1998). Although Taimela (1991) found in a Finnish sample that height was significantly associated with RT, height accounted for far less than 1% of the variance in RT.

The second explanation is that RT has increased because the sedentary lifestyle has become more common in Western countries. That is, the increase in RT reflects the fact that people have become more sluggish. Contradicting this explanation, Emery, Huppert, and Schein (1995) found in an adult British sample that the amount of walking people did was unrelated to their RT, and Jorm, Anstey, Christensen, and Rodgers (2004) found in an Australian sample that at three age levels (20–24, 40–44, and 60–64) activity level was unrelated to RT.

The third explanation is that RT has been increased by the buildup of neurotoxins in the environment. A great deal of research shows that RT is longer in people who have been exposed to heavy concentrations of neurotoxic substances, such as lead (e.g., Barth et al., 2002; Schwartz et al., 2001), chlordane (Kilburn & Thornton, 1995), and trichloroethylene and mercury (Liang, Sun, Sun, Chen & Li, 1993). We do not know whether RT is increased by low levels of neurotoxins, but we do know that exposure to neurotoxins in the everyday environment is widespread. Consider that a recent review listed 201 chemicals known to have neurological effects in humans and that most of them are in common use (Grandjean & Landrigan, 2006). Consider also that in the United States alone, for 2007 approximately 22,000 facilities reported 4.1 billion pounds of on-site and off-site disposal or releases of nearly 650 toxic chemicals (U.S. Environmental Protection Agency, 2007).

Although it is a plausible hypothesis that environmental neurotoxins have increased RT, it is difficult to make an exact prediction as to their effects over time. To do so would entail knowing for different study samples what neurotoxins they had been exposed to and at what concentrations and for what durations or having a comprehensive set of measurements of the levels of neurotoxins in the bodies of the study samples. Because we do not have this information for any of the comparison studies, this is yet another reason for not regressing RT on the year in which the study was conducted.

Another point to consider is that we should not assume that over the past century or so there have been only increases in exposure to environmental neurotoxins. Lead is a case in point. Until recent decades, exposure to lead was extensive, with the major sources of exposure coming from the lead that was added to gasoline and interior residential paint. Today the situation is much changed because the use of lead has come under tight government regulation in many countries. In the United States, for example, federal agencies, beginning in the 1970s, mandated reductions in the use of lead in gasoline and paint intended for residential use, and as environmental exposure to lead has been lowered, sharp declines have occurred in the mean blood lead levels of children living in the United States (Bellinger & Bellinger, 2006).

A final possible explanation for the increase in RT over time is based on the idea that the adult population has changed over time, in some ways for the worse. This explanation is based on two premises. One is that as the rates of infant and child death declined over the past 200 years (Riley, 2001), more people survived into adulthood who were in less than robust health. The other premise is based on research showing that people in less than robust health have longer RTs (Emery et al., 1995) and that life expectancy is inversely related to the length of RT (Deary & Der, 2005b; Shipley, Der, Taylor, & Deary, 2006). Thus, this explanation is that RT has increased over time because across birth cohorts there were increasing proportions of adults who were in less than robust health.

This explanation may seem counterintuitive because it is reasonable to believe that the health of the adult population has generally improved as the life span has increased over the past 200 years. However, longevity and robust health are not the same. This difference is illustrated by diabetes. Although advances in modern medicine have enabled many diabetics to remain alive and active, it is also the case that diabetics are at risk for poor circulation, vision loss, poor kidney functioning, infection, and nerve disorders (*Professional Guide to Diseases,* 2005). That diabetes also increases RT was shown in two recent studies (Anstey, Dear, Christensen, & Jorm, 2005; Pavlik, Hyman, & Doody, 2005). Not incidentally, the proportion of people ages 18–79 with diagnosed diabetes increased more than 100% between 1980 and 2007 in the United States (Centers for Disease and Control and Prevention, 2009).

Future research

It is not possible to say when RT began to increase and at what rate it changed. Nor is it possible to say what caused the increase in RT. However, two plausible causes for the increase in RT were proposed:

increased levels of environmental neurotoxins and declines in infant and child mortality rates. Both of these possibilities can be subjected to empirical test. With regard to the first possibility, it seems likely that as awareness about the deleterious effects of exposure to environmental neurotoxins increases, government controls over the production and use of these substances will become more stringent. If so, we should find a decrease in RT across birth cohorts as more and more such controls are put into place. With regard to the second possibility, RT should increase across birth cohorts in countries in which the infant and child mortality rates decrease. Of course, declines in exposure to environmental neurotoxins and declines in infant and child death rates are likely to go hand and hand, making it difficult to determine whether one or both of these secular changes are responsible for changes in RT over birth cohorts. However, it is possible that some countries will experience declines in infant and child death rates due to improvements in public health but at the same time experience increases in exposure to environmental neurotoxins due to increased use of neurotoxins in agriculture or manufacturing. These countries will be perfect laboratories in which to unravel the effects of the proposed causes for the increase in RT documented here.

As noted by an anonymous reviewer, a possible wildcard in making predictions for changes in RT across birth cohorts is that playing action video games has become widespread, and this might affect simple RT across birth cohorts. This may be less of a problem than what appears at first sight, because although action video games emphasize fast responding, they also require decision making and are therefore closer in format to choice RT tasks and go/no-go RT tasks than to simple RT tasks. On this point, the results obtained by Kida, Oda, and Matsumura (2005) are instructive. These investigators found that in high school and college-age males, the amount of experience playing baseball was positively correlated with speed of response on a go/no-go task but was unrelated to speed of response on a simple visual RT task. In baseball, players must learn to swing the bat at balls only in the strike zone. Thus, even though the bat must be swung quickly at the rapidly moving ball, the player must first decide whether to swing at the ball. Likewise, in many action video games the player must make super-quick decisions as to what actions to take, for example, to navigate an obstacle-filled course. So playing action video games also may not benefit simple RT.

Two recommendations for future research also emerge from the present study. One recommendation is to standardize the testing procedure for RT with respect to size, intensity, and duration of the stimulus and the number and pacing of the test trials. Although experimental psychologists have not shown much interest in standardizing their tests, there is reason to believe that in coming decades, we will be in a better position to make comparisons across studies of RT. The reason for hope in this regard is that RT tests have been incorporated into standardized neurobehavioral assessment batteries (e.g., Anger et al., 1993), batteries that are being used increasingly to assess, among other things, the effects of neurotoxins on human functioning.

A second recommendation is that extreme RTs must be handled in a uniform way. As seen in Table 2, whereas some investigators have chosen to ignore extreme RTs, other investigators have elected to delete them from the RT distribution. If RT distributions were symmetrical, removal of equal numbers of the longest and shortest RTs from the distribution would have no effect on the overall mean RT. But how do RTs distribute themselves? None of the comparison studies provide any information on this question. Gamberale, Iregren, and Kjellberg (1990) presented a frequency distribution of simple visual RTs obtained in a sample composed of 424 industrial workers; each worker was presented 160 test trials, for a total of 67,840 RTs. The frequency distribution is normal except for a slight positive skewing that probably reflects lapses in attention. If this frequency distribution is representative of other distributions for simple visual RT, retention of these very slow responses probably would have only a very small effect on the overall mean. An alternative approach for dealing with very long RTs is to use statistical models so as to screen out the RTs that presumably reflect inattention (Ratcliff, 1993; Heathcote, 1996). Yet another approach for dealing with the problem of very long RTs is to minimize their occurrence by using only a small number of test trials. This should not only decrease the frequency of very long RTs due to inattention but also should provide more accurate assessments of participants' optimal RT performance. The latter inference is based on the results of two studies (Gamberale et al., 1990; Olson, Gamberale, & Iregren, 1985), which show that RT increased across test trials within each of two test sessions.

Closing remarks

Two points should be made in closing. The first is that the present review suggests that performance has slowed over time for other psychomotor tasks as well. To test this expectation, the literature was searched for reviews in which performance on one or more psychomotor tasks was examined across birth cohorts. This search came up empty-handed. Whether the absence of such reviews reflects a paucity of relevant studies or the failure of psychologists to take an interest in secular trends for psychomotor performance is unclear. Whatever the reason, it is hoped that the present review will bring attention to the study of birth cohort changes in psychomotor abilities and other basic psychological functions.

The second point is that the study of birth cohort changes requires that psychologists take a very long time perspective in conducting their research. Although there has been a great deal of longitudinal research, this requires only the commitment to study people over decades. In contrast, to study secular changes requires the commitment to study people over generations. Clearly, research on birth cohort changes in human psychological functions will not be done unless institutions sponsor and nurture such undertakings. But as natural history museums have devoted large sums of money to the study the evolutionary history of humans, it would seem that these institutions (and perhaps others) should be interested in charting changes in human psychological functions over future generations.

NOTES

The author is grateful to Michael E. Doherty and Herbert M. Lefcourt for their comments on an earlier version of this article.

Address correspondence about this article to Irwin W. Silverman, Department of Psychology, Bowling Green State University, Bowling Green, OH 43403 (e-mail: silpsych@comcast.net).

1. These results raise the question of whether the RTs obtained in the comparison studies were longer than those obtained by Galton because, owing to a decrease in attention, there was a decline in performance over test trials in the comparison studies. That the number of test trials used in the comparison studies can explain the increase in RT is dubious because Gamberale et al. (1990) state that RT was on average 17 ms longer in the fifth than the first minute of testing. If performance declined gradually across test trials, as is likely, smaller differences in RT would have been reported between the fourth, third, and second minute of testing than the first minute of testing. It also should be noted that Gamberale et al. presented a total of 80 test trials per session, which exceeds the number of test trials in most comparison studies. Thus, although with fewer test trials the RTs in some of the comparison studies might have been shorter, it is likely that the decrease in RT would have probably been no more than 10 ms on average.

REFERENCES

Anger, W. K., Cassitto, M. G., Liang, Y.-X., Amador, R., Hooisma, J., Chrislip, D. W., et al. (1993). Comparison of performance from three continents on the WHO-recommended Neurobehavioral Core Test Battery (NCTB). *Environmental Research, 62,* 125–147.

Anstey, K. J., Dear, K., Christensen, H., & Jorm, A. F. (2005). Biomarkers, health, lifestyle, and demographic variables as correlates of reaction time performance in early, middle, and late adult. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology 58,* 5–21.

Barth, A., Schaffer, A. W., Osterode, W., Winker, R., Konnaris, C., Valioc, E., et al. (2002). Reduced cognitive abilities in lead-exposed men. *International Archives of Occupational and Environmental Health, 75,* 394–398.

Bellinger, D. C., & Bellinger, A. M. (2006). Childhood lead poisoning: The torturous path from science to policy. *Journal of Clinical Investigation, 116,* 853–857.

Boring, E. G. (1950). *A history of experimental psychology* (2nd ed.). New York: Appleton–Century–Crofts.

Brice, C. F., & Smith, A. P. (2002). Effects of caffeine on mood and performance: A study of realistic consumption. *Psychopharmacology, 164,* 188–192.

Centers for Disease Control and Prevention. (2009). *Data & trends*. Retrieved April 23, 2009, from http://www.cdc.gov/diabetes/statistics/incidence/fig2.htm

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences* (3rd ed.). Mahwah, NJ: Erlbaum.

Cole, T. J. (2000). Secular trends in growth. *Proceedings of the Nutrition Society, 9,* 317–324.

Deary, I. J., & Der, G. (2005a). Reaction time, age, and cognitive ability: Longitudinal findings from age 16 to 63 years in representative population samples. *Aging, Neuropsychology, and Cognition, 12,* 187–213.

Deary, I. J., & Der, G. (2005b). Reaction time explains IQ's association with death. *Psychological Science, 16,* 64–69.

Der, G., & Deary, I. J. (2006). Age and sex differences in reaction time in adulthood: Results from the United Kingdom Health Lifestyle Survey. *Psychology and Aging, 21,* 62–73.

Emery, C. F., Huppert, F. A., & Schein, R. L. (1995). Relationships among age, exercise, health, and cognitive function in a British sample. *Gerontologist, 35,* 378–385.

Flynn, J. R. (1984). The mean IQ of Americans: Massive gains 1932 to 1978. *Psychological Bulletin, 95,* 29–51.

Flynn, J. R. (1987). Massive IQ gains in 14 nations. What IQ tests really measure. *Psychological Bulletin, 1010,* 171–191.

Forbes, G. (1945). The effect of certain variables on visual and auditory reaction times. *Journal of Experimental Psychology, 35,* 153–162.

Galton, F. (1889). Instrument for determining reaction-time. *Report of the British Association for the Advancement of Science, 5,* 784–785.

Gamberale, F., Iregren, A., & Kjellberg, A. (1990). Computerized performance testing in neurotoxicology: Why, what, and whereto? In R. W. Russell, P. E. Flattau, & A. M. Pope (Eds.), *Behavioral measures of neurotoxicology: Report of a symposium* (pp. 359–394). Washington, DC: National Academy of Science.

Grandjean, P., & Landrigan, P. J. (2006). Developmental neurotoxicity of industrial chemicals. *Lancet, 368,* 2167–2178.

Heathcote, A. (1996). RTSYS: A DOS application for the analysis of reaction time data. *Research Methods, Instrumentation, & Computers, 28,* 427–445.

Jensen, A. R. (1980). *Bias in mental testing*. New York: Free Press.

Johnson, R. C., McClearn, G., Yuen, S., Nagosha, C. T., Abern, F. M., & Cole, R. E. (1985). Galton's data a century later. *American Psychologist, 40,* 875–892.

Jorm, A. F., Anstey, K. J., Christensen, H., & Rodgers, B. (2004). Gender differences in cognitive abilities: The mediating role of health state and health habits. *Intelligence, 32,* 7–32.

Kida, N., Oda, S., & Matsumura, M. (2005). Intensive baseball practice improves the go/no-go reaction time but not the simple one. *Cognitive Brain Research, 22,* 257–264.

Kilburn, K. H., & Thornton, J. C. (1995). Protracted neurotoxicity from chlordane sprayed to kill termites. *Environmental Health Perspectives, 103,* 690–694.

Kilburn, K. H., Thornton, J. C., & Hanscom, B. (1998). Population-based prediction equations for neurobehavioral tests. *Archives of Environmental Health, 53,* 257–261.

Koga, Y., & Morant, G. M. (1923). On the degree of association between reaction times in the case of different senses. *Biometrika, 15,* 346–372.

Krieg, E. F., Jr., Chrislip, D. W., Letz, R. E., Otto, D. A., Crespo, C. J., Brightwell, W. S., et al. (2001). Neurobehavioral test performance in the third National Health and Nutrition Examination Survey. *Neurotoxicology and Teratology, 23,* 569–589.

Ladd, G. T., & Woodworth, R. S. (1911). *Physiological psychology*. New York: Scribner.

Lefcourt, H. M., & Siegel, J. M. (1970). Reaction time behaviour as a function of internal–external control of reinforcement and control of test administration. *Canadian Journal of Behavioural Sciences, 2,* 253–266.

Liang, Y.-X., Sun, R.-K., Sun, Y., Chen, Z., & Li, L.-H. (1993). Psychological effects of low exposure to mercury vapor: Application of a computer-administered neurobehavioral evaluation system. *Environmental Research, 60,* 320–327.

Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. Thousand Oaks, CA: Sage.

Macey, T., Plomin, R., & McClearn, G. (1979, June). *Evidence for the reliability of data from Galton's anthropometric laboratory*. Paper presented at the meeting of the Behavior Genetics Association, Middletown, CT.

Olson, B. A., Gamberale, F., & Iregren, A. (1985). Coexposure to toluene and p-xylene in man: Central nervous functions. *British Journal of Industrial Medicine, 42,* 117–122.

Pavlik, V. N., Hyman, D. J., & Doody, R. (2005). Cardiovascular risk factors and cognitive functions in adults 30–59 years of age (NHANES III). *Neuroepidemiology, 24,* 42–50.

Professional guide to diseases (8th ed.). (2005). Ambler, PA: Lippincott Williams & Wilkins.

Ratcliff, R. (1993). Methods for dealing with reaction time outliers. *Psychological Bulletin, 114,* 510–537.

Reed, T. E., Vernon, P. A., & Johnson, A. M. (2004). Sex difference in brain nerve conduction velocity in normal humans. *Neuropsychologica, 42,* 1709–1714.

Riley, J. C. (2001). *Rising life expectancy: A global history*. New York: Cambridge University Press.

Robinson, E. S. (1934). Work of the integrated organism. In C. Murchison (Ed.), *A handbook of general experimental psychology* (pp. 571–650). Worcester, MA: Clark University Press.

Rose, K., Smith, W., Morgan, I., & Mitchell, P. (2001). The increasing incidence of myopia: Implications for Australia. *Clinical & Experimental Ophthalmology, 29,* 116–120.

Schwartz, B. S., Lee, B.-K., Le, G.-S., Stewart, W. F., Lee, S.-S., & Hwang, K.-Y. (2001). Associations of blood lead, dimercaptosuccinate acid–chelatable lead, and tibia lead with neurobehavioral test scores in South Korean lead workers. *American Journal of Epidemiology, 153,* 453–464.

Seashore, R. H., Starmann, R., Kendall, W. E., & Helmick, J. S. (1941). Group factors in simple and discrimination reaction times. *Journal of Experimental Psychology, 29,* 346–394.

Shipley, B. A., Der, G., Taylor, M. D., & Deary, I. J. (2006). Cognition and all-cause mortality across the entire adult age range: Health and lifestyle survey. *Psychosomatic Medicine, 68,* 17–24.

Silverman, I. W. (2006). Sex differences in simple visual reaction time: A historical meta-analysis. *Sex Roles, 54,* 57–68.

Smith, A., Sturgess, W., Rich, N., Brice, C., Collison, C., Bailey, J., et al. (1999). The effects of idazoxan on reaction times, eye movements and the mood of healthy volunteers and patients with upper respiratory tract illness. *Journal of Psychopharmacology, 13,* 148–151.

Storfer, M. (1999). Myopia, intelligence, and the expanding human neocortex: Behavioral influences and evolutionary implications. *International Journal of Neuroscience, 98,* 153–276.

Taimela, S. (1991). Factors affecting reaction-time testing and the interpretation of results. *Perceptual and Motor Skills, 73,* 1195–1202.

Taimela, S., Kujala, U. M., & Osterman, K. (1991). The relation of low grade mental ability to fractures in young men. *International Orthopedics, 15,* 75–77.

Teichner, W. H. (1954). Recent studies in simple reaction time. *Psychological Bulletin, 51,* 128–149.

Tomkinson, G. R., Léger, L. A., Olds, T. S., & Cazorla, G. (2003). Secular trends in the performance of children and adolescents (1980–2000): An analysis of 55 studies of the 20 m shuttle run test in 11 countries. *Sports Medicine, 33,* 285–300.

U.S. Environmental Protection Agency. (2007). *2007 TRI public data release.* Retrieved October 2, 2009, from http://www.epa.gov/tri/tridata/tri07/index.htm

Wilkinson, R. T., & Allison, S. (1989). Age and simple reaction time: Decade differences for 5,324 subjects. *Journal of Gerontology, 44,* 29–35.

Woodworth, R. S., & Schlosberg, H. (1954). *Experimental psychology* (Rev. ed.). New York: Henry Holt.

World Health Organization. (2008). *Obesity and overweight.* Retrieved July 1, 2008, from http://www.who.int/mediacentre/factsheets/fs311/en/print.html