

Timeless demonstrations of Parkinson's first law

LAURA A. BRANNON

University of Oklahoma, Norman, Oklahoma

PAUL J. HERSHBERGER

Veterans Affairs Medical Center, Dayton, Ohio

and

TIMOTHY C. BROCK

Ohio State University, Columbus, Ohio

"Work expands so as to fill the time available for its completion," Parkinson's law, is an explanation classic that has survived without an artifact-free demonstration at the individual level. To evaluate Parkinson's law, undergraduate subjects expected to judge four sets of photos of faces with reference to a subjective criterion. The experimental subjects, who were told that the fourth set was canceled before they began work on the third set, dallied on the third set; that is, as compared with controls, they prolonged work. The cancellation-dalliance effect was reobtained in two exact replications. It was obtained again in a fourth study, a conceptual replication wherein subjects processed negatively toned phrases against an objective criterion. The generalizability of the effect and explanations for it are discussed.

The temporal elasticity of a set of activities, the degree to which the activities take more or less time as a function of the circumstances under which they are carried out, is the central idea of Parkinson's notorious first law. . . . It seems to us an important idea, one that deserves more study than it has thus far received. (McGrath & Kelly, 1986, p. 124)

"Work expands so as to fill the time available for its completion" became an explanation classic soon after its promulgation as a "law" by Parkinson (1957). Four decades later, the familiar aphorism continues to be evoked to account for a plethora of human and organizational inefficiencies. However, although Parkinson's formulation has become a household expression, its popularity never elicited a convincing empirical demonstration.

Parkinson's First Law: Whence and Whither

When Parkinson's (1957) mathematical formula for his "law" is examined, its satirical character is apparent; although tongue-in-cheek, it expressed sincere revulsion at metastasizing bureaucracies.¹ Parkinson later examined additional data and concluded that he should be regarded "as a true prophet" (Parkinson, 1980, p. 23) because "work *does* expand so as to fill the time available" (ibid. p. 35, Parkinson's italics). In 1980, as in his initial 1957 statement, the evidence adduced was increases in *aggregate* numbers of staff members rather than a direct

study of the persons who were the presumed perpetrators of work expansion.

Mixed Endorsement by Scholars

Parkinson's law is widely taught in a variety of disciplines that deal with behavior in work settings. While the quality of the evidence for the "law" appeared to be beyond question in most sociology textbooks (Giddens, 1989, p. 277; Kornblum, 1988, pp. 362-363; Macionis, 1991, pp. 187-188; Thio, 1991, pp. 87-88), in other textbooks (e.g., Conklin, 1984, p. 163) there is acknowledgment of failures to confirm (e.g., Reimann, 1979). Similarly, in many organizational behavior and management texts, the validity of the "law" is accepted, and it plays an important explanatory role (Carvell, 1980, p. 419; Hersey & Blanchard, 1988, p. 128; Koontz & Weirich, 1990, p. 12; Luthans, 1981, pp. 446-447; Miner, 1980, pp. 170-177; and Robbins, 1990, pp. 162-167). However, neither accepting nor doubting scholars have been able to rely on experimental tests of individuals' susceptibility to excess time effects.

Aronson's Experimental Explorations "Beyond Parkinson's Law"

In quest of compelling evidence for Parkinson's law, the early results of Aronson and Gerard (1966), Aronson and Landy (1967), and Landy, McCue, and Aronson (1969) were examined. Unfortunately, the Aronson studies were inconclusive because, rather than demonstrating the excess time effect, it was simply taken for granted and, also, because time was artifactually emphasized for the research subjects. In both the 1966 and 1967 experiments, Aronson allowed research subjects too much time

The comments of Joe McGrath, Janice Kelley, and Craig J. Russell on earlier drafts are acknowledged with gratitude. Our indebtedness to several anonymous reviewers for their comments on an earlier draft is also acknowledged. Correspondence concerning this article should be addressed to L. A. Brannon, Department of Psychology, University of Oklahoma, Norman, OK 73019 (e-mail: lbrannon@ou.edu).

to perform a task. Subsequently, when presented with a similar task and allowed to work at their own pace, subjects who initially were allowed excess time required more time to complete the subsequent task. The subject was seated at a table with a timer and asked to prepare a 2-min talk (with materials supplied by the experimenter); then an interruption occurred, the critical manipulation of excess time. A secretary "barged in," asked the experimenter to help another experimenter, and announced that it would take 15 min (or 5 min). The experimenter "gave in resignedly, 'Only 15 (5) minutes, eh?'" When the experimenter returned, the subject was asked to work on preparing another speech (Aronson & Gerard, 1966), and time on this second task was the principal dependent variable.

The data did show that subjects who were allowed more time for the first task took more time on the second task. However, the phenomenon of excess time on the *first* task was not examined and the procedure was replete with time cues: the subjects faced the timer on the table, and the interruption of the experimenter was attended by emphatic time stipulations—for example, "only fifteen minutes; only five minutes."

The 1967 (Aronson & Landy) replication employed a different task and included a further manipulation, with the second task being similar to or different from the first task. However, when the initial task was explained to the subjects, the interruption manipulation had the same prominent time-cuing features. Subjects who were allowed 5 min (and knew they were allowed just 5 min) utilized an average of 288 sec, while subjects who were allowed 15 min (and knew they had 15 min) spent an average of 640 sec [$F = 66.94, p < .001$; effect size was calculated as $d = 2.02$; Johnson, 1993].

Unfortunately, because of the aforementioned pervasive time cuing, this possible demonstration of the Parkinson effect could not be convincing. Importantly, Aronson and his co-workers acknowledged this ambiguity: "differential time norms may have been conveyed by E as to what constituted an appropriate time to spend on a given task" (Landy et al., 1969, pp. 236–237). In their final experiment, Aronson and his coexperimenters nevertheless continued the time cuing and even increased it. An interruption with time announcements was used again. Furthermore, the subject was confronted by a confederate who sat very close to the subject (allegedly because all the other cubicles were in use) and repeatedly asked the subject to hurry. Of course, subjects told to hurry did take less time than subjects in a control condition in which they were not badgered to hurry. In sum, the Aronson studies may have gone "beyond" Parkinson's law but they did not demonstrate the law itself.²

In this first stringent test of Parkinson's prediction, we sought to eliminate time demands and time cues while probing for dawdling on their current work by subjects who had less future work to do than they expected. Our first three studies used a photo-ranking procedure, sim-

ilar to that used by Aronson and Landy (1967; Landy et al., 1969); it required subjects to look at photographs and rank-order the people depicted in terms of various criteria, such as *subjective* impression of intelligence. Use of a similar task was considered critical to the cumulativeness of the present investigations with Aronson's earlier findings. Our final experiment, Study 4, was a conceptual replication in which subjects processed negatively toned phrases against an *objective* criterion.

THREE STUDIES OF TIME SPENT JUDGING SETS OF PHOTOS

Method

Overview

Subjects individually followed a procedure in which they expected four trials but actually completed only three. A trial consisted of judging photos of faces with reference to a criterion. The experimental subjects were told about the "cancellation" of the fourth trial just before they began work on the third, the test, trial; they thus had "extra" time. In contrast, the control subjects expected a fourth trial as they worked on the third trial, the test trial. The primary dependent measure was the change in time taken to complete the third (test) trial as compared with the time taken to complete the second (baseline) trial. Oral instructions were used in Study 1, whereas written instructions were used in Studies 2 and 3.

Subjects

The subjects were undergraduate students, male and female, from introductory psychology classes at a Midwestern university, who participated in the study in partial fulfillment of a course requirement: $N = 22$ (Study 1), $N = 47$ (Study 2), and $N = 109$ (Study 3). Importantly, all subjects from this pool expected to spend up to an hour in each required research participation.

Materials

The stimulus materials consisted of four sets of 10 pictures, 20 of males and 20 of females. The portrait pictures were taken from newspaper wedding pictures so that all persons would be dressed in formal attire. This source for the pictures also meant that educational and occupational information about the persons depicted was plausibly obtainable from the accompanying announcements. The pictures were trimmed so that only faces of the brides and of the grooms were visible.

Procedure

Individually tested subjects sat facing a wall, so their field of vision excluded both the experimenter and the timekeeper. The timekeeper had no interaction of any kind at any time with the subject.

The subjects were given instructions which indicated that the experiment was attempting to determine whether or not it was possible to accurately estimate the educational and occupational levels of persons by looking at pictures of the persons' faces. The instructions further indicated that the experiment would consist of two equal parts and that each part would involve ranking two sets of pictures. The text of the instructions was as follows:

This experiment is part of an exploratory study in which we are trying to determine whether it is possible to guess the education and occupational level of persons simply from looking at their pictures. This is a skill that could be very important in police work, for persons in service occupations, and for fundraisers. Our procedure is quite straightforward. The experiment has two parts. In Part 1 of the experiment, you will rank pictures of 10 men according to your best guess of their education and occupational level and then 10 women based on the same cri-

teria. Part II of the experiment is identical to Part I, except that you will be using different pictures.

Rank the pictures from lowest education/occupational level to the highest. For example, someone whom you believe has a high school diploma but no further formal education would be ranked before someone whom you believe has completed graduate school. In the same way, an assembly line worker would be ranked before a high level corporate executive.

At this point the timekeeper began surreptitious timing. The timing was stopped when the subject indicated that he or she was finished with the set of pictures. During this practice/warm-up trial, questions were allowed from the subjects and were answered by the experimenter. The practice trial enabled standard administration of the next trial, the baseline trial.

Trial cancellation manipulation. Control-condition subjects completed the two picture sets (the practice/warm-up and the baseline trials) in Part I, and were then told that Part I was now completed and that the procedure in Part II would be exactly the same. The complete text of the instructions was: "You have now completed Part I of the experiment. The identical procedure will be followed in Part II of the experiment."

After completing the first set of pictures in Part II (the test trial), the control-condition subjects were told that there would not be another set to rank (no fourth trial). Thus, the control-condition subjects completed all three trials with no basis for suspecting that the fourth trial would not occur. An additional control group ($N = 14$) in Study 3 was told at the outset that there would be only three trials.

Experimental-condition subjects also completed the two picture sets in Part I (the practice/warm-up and the baseline trials), but these subjects were then told that one of the sets of pictures for Part II of the experiment had been canceled. Therefore, they would have only one picture set to rank in Part II. The complete text of the instructions was: "I'm unable to find the pictures of the women for Part II of the experiment. Instead of ranking both men and women in Part II of the experiment, you'll only rank 10 men. Here are the 10 pictures, and again, let me know when you are finished."

Finally, the subjects were fully debriefed: the actual purpose was disclosed, Parkinson's law was defined and illustrated, and the cancellation manipulation was explained.

Dependent measures. The timekeeper surreptitiously scored, by stopwatch, the beginning and end of the subject's use of the photos. The principal dependent variable was the change in the amount of time spent on the test trial (the picture ranking in Part II) as compared with the baseline trial (the second picture ranking in Part I).

Probing task perceptions in Studies 2 and 3. After ranking the one set of pictures in Part II (the test trial), the subjects were also asked to answer five posttask questions:

1. Did you find the sets of pictures to be of equal interest? If no, which more interesting?
2. Did you find the sets of pictures to be of equal difficulty? If no, which more difficult?
3. Estimate the actual time you spent on each of the sets, in minutes and seconds
4. Did you feel you were being evaluated in any way beyond what was explained at the beginning of the experiment?

Experimental subjects were further asked about the thoughts that went through their minds when they were informed that the experimenter had misplaced one of the sets of pictures for Part II; they were also asked, "How did this affect the amount of time you spent on the third set?"

Results

All subjects readily completed their photo-rankings according to the instructions. No subject's data have

been omitted for any reason. The principal results for the three studies are presented in the top, middle, and bottom panels of Figure 1. The left side of each panel shows mean speed in seconds for the baseline and test trials for the control condition (no cancellation, $N = 11$ in Study 1; $N = 22$ in Study 2; and $Ns = 46$ and 14 in Study 3); the right side of each panel shows the same data for the experimental conditions (fourth trial canceled, $N = 11$ in Study 1; $N = 25$ in Study 2; $N = 49$ in Study 3). Gender of subject did not enter into any reliable main or interaction effects for the three studies. Gender is therefore not further discussed.³ The principal dependent variable was the change in the amount of time spent on the test trial (the picture ranking in Part II) compared with the baseline trial (the second picture ranking in Part I). The propensity of a subject to be quick or slow on the test trial was examined by a 2 (condition: experimental vs. control) \times 2 (trial: baseline vs. test) analysis of variance on the number of seconds required to complete the task, with repeated measures on the second factor. There was a significant condition \times trial interaction [$F(1,20) = 5.30, p < .04$, in Study 1 and $F(1,45) = 5.87, p < .02$, in Study 2]. Study 3 included a second control condition (expect three trials). A similar 3 (condition: experimental, control 1, control 2) \times 2 (trial: baseline vs. test) analysis with repeated measures on trial yielded a borderline significant condition \times trial interaction [$F(2,106) = 2.98, p = .055$].

As shown in Figure 1, the control subjects worked through the sets of pictures more quickly on the test trial than on the baseline trial. The times were also faster for the baseline trial than for the practice/warm-up trial (not shown). In contrast, while the experimental subjects also worked faster on the baseline trial than on the practice trial, these subjects tended to work more slowly on the test trial. In comparison with the control subjects, the experimental subjects dallied⁴ over the first set of pictures in Part II: work appeared to expand to fill the time available for its completion, as Parkinson proposed.

The parallel pattern for the second Study 3 control, expect three trials (bottom panel of Figure 1), showed that the cancellation-dalliance effect did not depend on comparison with another group (the primary control) that was also expecting a fourth trial. The comparison between the expect-three-trials group (control 2) and the experimental group also appeared to rule out the hypothesis that the latter worked relatively more slowly than the control group in order to prolong their enjoyment of the task. If that had been the case, the expect-three-trials subjects might also have worked more slowly to "savor" the final set of photos. An additional manipulation of expected trials is reported in Study 4 below.

Effect sizes, ds (Johnson, 1993) for Studies 1, 2, and 3 were .94, .70, and .46, respectively. These medium-to-high effect sizes were obtained with no explicit time cuing. Although robust, they were lower than the 2.02 calculated for the Aronson and Landy (1967) time-cued demonstration of the effect. However, our data and the

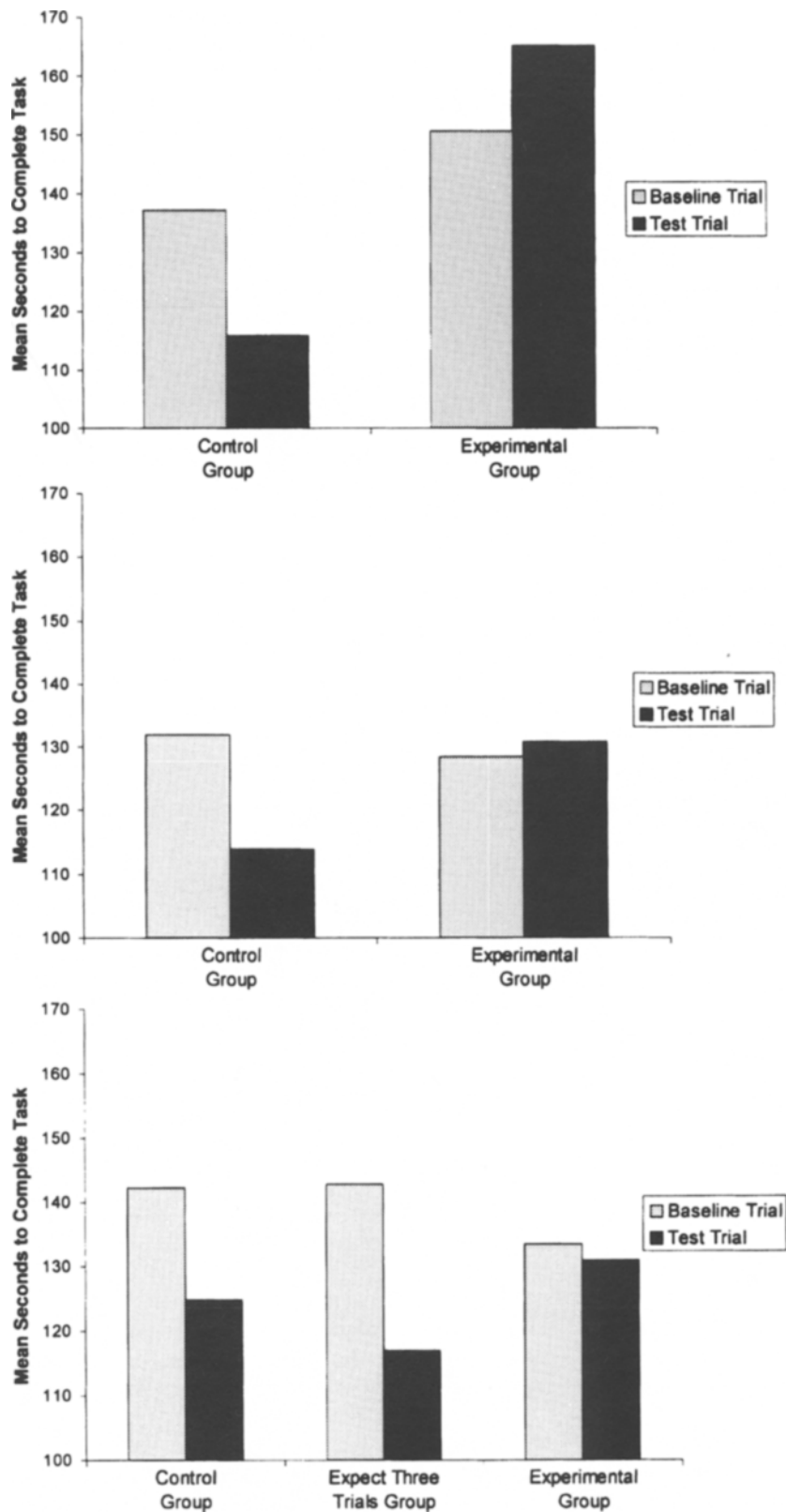


Figure 1. Effects of trial cancellation on work speed in seconds in Study 1, top panel; Study 2, middle panel; and Study 3, bottom panel.

Aronson data, taken together, suggest that the Parkinson effect is not only quite real but that it may be robust in real-life work settings because such settings do commonly include time cues and time-related accessories.

Task Perceptions (Studies 2 and 3)

In addition to using written instructions, Studies 2 and 3 probed subjects' perceptions with virtually the same results for both studies. Subjects' estimates of the time spent on each set of photos were entered into an index, namely, the difference between actual time and estimated time divided by the actual time taken. Whether verbal estimates were affected by condition was determined by 2 (condition: experimental vs. control) \times 2 (trial: baseline vs. test) analyses of variance, with repeated measures on trial. The interaction *F*s for these analyses for both Study 2 and Study 3 were below unity. When alternative indices were calculated, such as the simple difference between actual and estimated time and the ratio of each person's verbal estimate to his or her actual time, main and interaction effects for condition and trial again yielded *F*s < unity. Hence, there was no evidence that cancellation of the fourth trial caused experimental subjects to somehow arrive at more accurate recollection of their work durations on the test trial.

To further test subjects' sensitivity to dallying, *t* tests were conducted comparing the verbal estimates of third-trial duration made by subjects in the experimental and control conditions. In both Study 2 and Study 3, these differences yielded *p*s > .20. An apparent tendency for experimental subjects to report longer durations fell well short of conventional criteria for reliability.

There were no differences between experimental and control subjects in their perceptions of the interest and difficulty of the sets of pictures. Fewer than 5% of the subjects commented about being evaluated, and the only comment made by more than 1 subject was to surmise that the experiment dealt with comparing how males and females made such judgments.

Finally, subjects in the experimental conditions were reminded that the experimenter had misplaced pictures for Part II and were asked, "How did this affect the amount of time you spent on the third set?" Most subjects had no comment. The most important observation of these comments was that none of the subjects mentioned "tried harder," "tried to be more accurate," "spent more time on the pictures," or the like.

STUDY 4

Conceptual Replication With Negative Stimuli and Objective Measurement of Performance

A fourth experiment addressed a number of questions that were not resolved by the first three studies. Was the Parkinson effect limited to seemingly neutral stimuli? Would people prolong work that was inherently unpleas-

ant? Could the effect be observed for tasks in which work quality had an objective criterion? Was the effect proportional to the amount of further work that had been canceled? The first three studies deliberately used a task, ranking photos, that was quite similar to that used originally by Aronson (e.g., Aronson & Landy, 1967). However, that task did not allow for observation of dawdling in an activity that was somewhat unpleasant, and it did not allow for objective measurement of performance or for assessment of the possible impact of the magnitude of the canceled activity on amount of dalliance. If dalliance was related to the magnitude of the cancellation, the condition \times trial interactions of Studies 1, 2, and 3 might be qualified by a higher order interaction with number of canceled trials. Study 4 therefore utilized an objectively scorable and somewhat unpleasant task, and amount of cancellation was manipulated. Finally, although estimated durations were not affected by the treatments in Studies 1, 2, and 3, these estimates might have been insensitive measures of actual motivation to try harder. Consequently, we included new measures of subjects' impressions of the task and of their effort (Bryan & Locke, 1967).

Method

Overview

As in Studies 1, 2, and 3, subjects individually followed a procedure in which they expected 4 (or 12) trials but actually completed only 3. A trial consisted of counting the number of letters in a set of 25 phrases (see materials below). Cancellation instructions were read to the subjects by the experimenter, but other instructions were printed.

Subjects

The subjects were 86 undergraduate students, male and female, from introductory psychology classes at a Midwestern university, who participated in the study in partial fulfillment of a course requirement. As in Studies 1, 2, and 3, all subjects from this pool expected to spend up to an hour in each required research participation.

Materials

The stimulus materials consisted of three sets of 25 negative two-word phrases—one set for each trial. For example, in the test trial, some of the phrases were "suffocating air," "poisoned water," "gloomy day," "menacing dogs," "parking penalty," "vicious gossip," "criminal threat," "foul odors," and so forth. The word pairs were printed with variable spaces between them above a scale with the numbers from 8 to 28. For example, for "DISTRESSED CHILDREN" the correct letter-counting score was 18.

Procedure

Individually tested subjects, who were unaware that they were being timed, sat facing a wall in a cubicle so that their field of vision excluded both the experimenter and the timekeeper. The sur-reptitious timekeeper had no interaction of any kind at any time with the subject.

The subjects were given instructions which indicated that the experiment was attempting to determine how well reading and counting could be done when they were done simultaneously. The text of

the instructions was as follows with the expect-12-trials wording in brackets:

Reading and Enumeration Study:

The experiment today deals with two human processes, reading text and counting objects, and how well these basic processes can be done when they are done together. These basic processes, reading and counting, are essential to many kinds of work and often occur together in everyday life. The experiment has two parts. In the first section of Part I of the experiment, you will read and count the letters in a set of 25 phrases as well as you can. In the second section of Part I you will again read and count the letters in another set of 25 phrases. The two [ten] sections of Part II of the experiment are identical to Part I, except that you will be using different sets of phrases. Therefore, after Part I, there will be two [ten] more sets of phrases with 25 phrases in each set. In this study, read a PHRASE and then circle the number that represents the number of letters in the entire PHRASE.

A worked example followed.

The counting task was rendered nontrivial by varying the number of spaces between the words in the pair: it was not possible for a subject to correctly count by simply learning the counts that corresponded to the positions of the last letter in the second word.

At this point, the subject began work, and the work was timed. The timing was stopped when the subject indicated that he or she was finished with the first set (practice/warm-up) of phrases (word pairs). During the warm-up set, questions from subjects were allowed and answered by the experimenter. Then the subject was given the second set (baseline) of phrases to score and timing resumed.

Trial cancellation manipulation and debriefing. With respect to the third (test) set, the cancellation procedures for the succeeding set(s) and the debriefing for control and experimental subjects were similar to those of Studies 1, 2, and 3, except that subjects were told that the experimenter had just learned that the remaining sets had been canceled (rather than misplaced) and the number of remaining sets was 10 (rather than 1) in the expect-12-trials condition. The subjects' total work time on the baseline set of phrases averaged 199 sec ($SD = 30$ sec). Therefore, even subjects in the expect-12-trials condition could complete the task in approximately 40 min (much less than the 60 min allowed for a unit of research participation). No subject would have felt "hurried" (by the nature of the task or the number of anticipated trials) to finish.

Objective measures of performance and concentration/effort: Letter counting and surprise recall. As an objective measure of performance, we scored correct letter counting. In addition, in order to assess subject concentration/effort, we measured subjects' ability to recall one of the words from the test-trial word pairs when either the first or the second word was presented. Presumably, greater concentration/effort during the letter-counting task would result in greater recall of the word pairs on a surprise recall test.

Task perceptions. Study 4 subjects used a 5-point Likert scale to rate how pleasant and how interesting it was to read the phrases in each set. A self-assessment of work speed was adapted verbatim from Bryan and Locke (1967, p. 266):

- During the third set, the last set, I was trying to:
- ___ Work faster than on the first two sets
 - ___ Work at about the same pace as on the first two sets
 - ___ Work a little slower than on the first two sets
 - ___ Work as fast as possible
 - ___ Work quickly, but not as fast as possible
 - ___ Work with as little effort as possible
 - ___ Other.

Finally, we included Likert items to measure the extent to which the subjects were trying "harder to circle the correct number of letters in the last set of words" and the extent to which they were "hurrying to complete the experiment."

Results

All subjects were given as much time as needed to count characters in the three sets of word pairs; thus, the subjects achieved perfect counting scores on the third set. No subject data were deleted for any reason. Sex of a subject did not enter into any reliable main or interaction effects in Study 4 and gender is not further discussed.

Subjects found reading the phrases to be moderately unpleasant. The Likert-scale points were labeled 1 = *very unpleasant*, 2 = *unpleasant*, 3 = *neither pleasant nor unpleasant*, 4 = *pleasant*, and 5 = *very pleasant*. Overall means ($N = 86$) for the three word sets were 2.7, 2.5, and 2.4, respectively. The pleasantness ratings were unaffected by condition (experimental vs. control) or expected trials (4 vs. 12) or the interaction of these factors. Similar Likert ratings of "interesting" were slightly below the midpoint, 2.9, for the three sets of phrases and were unaffected by the independent variables. Therefore, whatever dalliance was observed was not attributable to different levels of task enjoyment or of task interest.

The principal dependent variable was the change in the amount of time spent on the test trial (counting the number of characters in the third set of phrases) compared with the baseline trial (counting the second set of phrases). The graphs of Figure 2 show the mean speeds in seconds for the baseline and test trials for control ($N = 21$) and experimental ($N = 24$) groups under expect-4 trials (leftmost) and for the control ($N = 18$) and experimental ($N = 23$) groups under expect-12 trials (rightmost). The propensity of a subject to be quick or slow on the test trial was examined by a 2 (condition: experimental vs. control) \times 2 (expected trials: expect 4 vs. expect 12) \times 2 (trial: baseline vs. test) analysis of variance on the number of seconds required to complete the task, with repeated measures on the third factor. There was a significant condition \times trial interaction [$F(1,82) = 75.69, p < .0001$]; this interaction reproduced the same pattern that was observed in Studies 1, 2, and 3. The effect size (Johnson, 1993) was $d = 1.87$, a high outcome comparable with the $d = 2.02$ calculated (above) for Aronson's data (Aronson & Landy, 1967).

In addition, there was a condition \times expected trials \times trial interaction [$F(1,82) = 10.02, p < .002, d = .68$]. Figure 2 shows that the higher order interaction was attributable to higher dalliance in response to greater cancellation. The experimental subjects' mean work duration on the test trial under expect-12 trials, 233, was reliably higher than their mean duration under expect-4 trials, 211 [$t(45) = 2.2, p < .03, d = .63$]. Thus, dalliance here appeared to be proportional to the magnitude of the canceled work.

Objective Measures of Performance and Concentration/Effort: Letter Counting and Surprise Recall

As noted previously, given the relative simplicity of the letter-counting task, all subjects, regardless of condition (experimental vs. control), correctly counted the

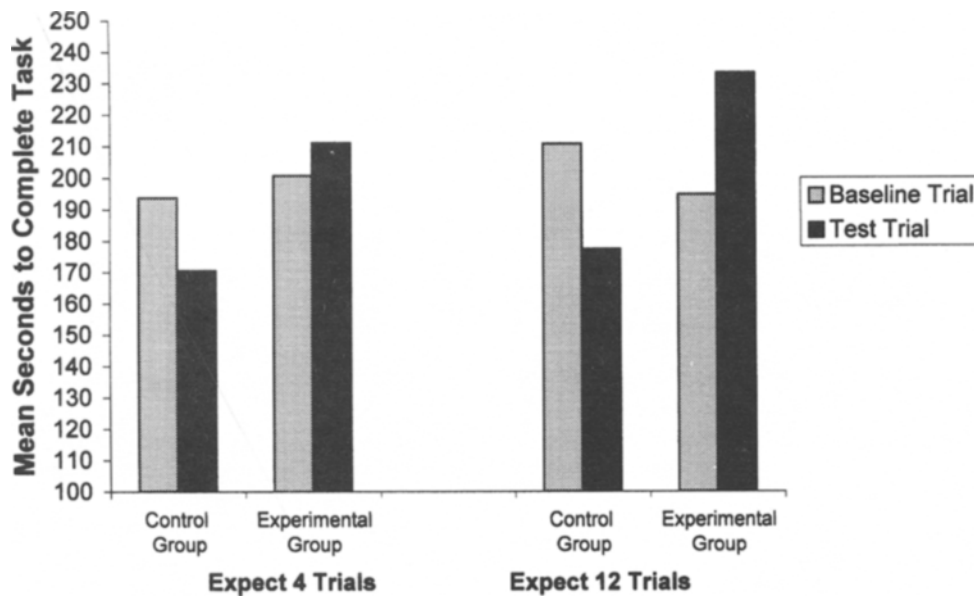


Figure 2. Effects of trial cancellation on work speed in seconds in Study 4 (expect 4 trials, left panels; expect 12 trials, right panels).

number of letters in each set. Dalliance in the experimental condition was not associated with better performance on the counting task.

Recall of word pairs was difficult for the subjects, in part because they did not expect a recall test. When the third set of words was readministered with one word in each set missing, overall correct mean recall of the missing words was 5.1 (obtained range = 0 to 14; perfect recall = 25). Between-subjects analyses showed that recall performance was not affected by condition (experimental vs. control) or by expected trials (4 vs. 12) or by their interaction. Moreover, recall and third-trial performance time were only slightly correlated ($r = .13$, n.s.), and the within-condition correlations in the experimental and control groups did not differ from each other. The obtained range of scores (0 to 14) indicated that the absence of effects of the independent variables was not attributable to a restriction of range. Dalliance was not associated with greater concentration/effort.

Task Perceptions

To measure subjects' assessment of their own motivation, we adapted a question from Bryan and Locke (1967, p. 266). With respect to the test trial, 69% of the control subjects and 72% of the experimental subjects used the response alternatives "work faster than on the first two sets" and "work at about the same pace as on the first two sets." Chi-square analyses for these frequencies and frequencies for all response alternatives (see Procedure above) yielded no reliable differences between the control and experimental conditions. With respect to the test trial question, only 2 experimental subjects and 2 control subjects said "worked a little slower than on the

first two sets." To the extent that this question was sensitive to subjects' motivation, the results indicated that experimental subjects were not aware of their actual slowing down relative to the control subjects.

We also asked subjects to what extent they were "trying harder to circle the correct number of letters on the last set of words, the third set, than on the previous set of words." The 5-point Likert item was anchored with 1 = "I tried much harder on the third set than the second set" and 5 = "I tried much harder on the second set than on the third set." The means for the control group were, under expect-4 trials, $M = 2.6$ and, under expect-12 trials, $M = 3.0$; the corresponding means for the experimental group were 2.5 and 3.0. The only reliable effect was that subjects under expect-4 trials reported trying harder on the last set [$F(1,82) = 5.54$, $p < .02$]. Condition itself (experimental vs. control) did not affect self-report of trying harder on the third than on the second set. Once again, cancellation, the Parkinson treatment, did not influence reported motivation. The final Likert item, "To what extent were you hurrying to complete the experiment," had five anchors: "A lot," "Somewhat," "A little," "Slightly," and "Not at all." The overall mean fell between "somewhat" and "a little," but there were no effects due to treatments. In sum, there was no tendency for experimental subjects, as compared with controls, to report trying harder on the third set or to report hurrying less to complete the experiment.

GENERAL DISCUSSION

The present studies have consistently demonstrated Parkinson's "law" under conditions free of the explicit

time-cuing and time demands that confounded previous research by Aronson and his co-workers (Aronson & Gerard, 1966; Aronson & Landy, 1967; Landy et al., 1969) and by Bryan and Locke (1966). While the Aronson studies appeared to demonstrate that taking excess time instigated further taking of excess time, the basic Parkinson phenomenon was never unambiguously demonstrated. In contrast to subjects who were externally paced by investigator-provided cues, the present subjects paced themselves: in this context, an unexpected cancellation instigated dalliance. Furthermore, the three-way interaction of Study 4, condition \times expected trials \times trial, suggested that magnitude of dalliance might be affected by the magnitude of the canceled work. In other words, the more unexpected time people have, the more they dally.

The present experiments comprised two simple replicable Parkinson-effect paradigms, one with seemingly neutral stimuli and a subjective task and the other with moderately unpleasant stimuli and an objective task. In both paradigms, next-task cancellation increased work time on the focal task even though all subjects knew that research participation required up to an hour. Expansion of time by the experimental subjects in Study 4 did not result in better work (accuracy in counting characters in word pairs was unaffected) or more concentration/effort (there was no better recall of the word pairs on a surprise recall test).

To make our findings cumulative with Aronson's (e.g., Aronson & Landy, 1967), we used a similar task in Studies 1, 2, and 3 in which performance could not be evaluated objectively. But, even if subjects who dallied on the task did not actually do a better job, is it possible that doing better work became their purpose? In Study 4, dalliance did not improve task performance, but was better performance an aim nonetheless?

The postexperimental questionnaires had several items which might have revealed that experimental subjects, in contrast to the controls, were trying harder, aiming more for accuracy, and so forth. As shown in Studies 1, 2, and 3, on none of these scaled and open-ended questions did the experimental subjects differ from the controls. Furthermore, since there was ample time left over, it is doubtful that control subjects were hurrying beyond what was reasonable in the situation in order to "fit in" extra trials. In Study 3, control subjects accelerated their work from baseline to test trial in parallel with a second control group, which had expected three trials from the outset (bottom panel of Figure 1).

New probes in Study 4 disclosed no differential effort or perception of effort by subjects who experienced next trial cancellation as compared with those who completed Trial 3 without knowing it was the last trial. Experimental and control subjects appeared to have the same task goals. Thus, the longer durations on the test trial for the experimental subjects cannot be readily attributed to ei-

ther conscious "hurrying" by control subjects or to conscious "lingering" or extra effort by experimental subjects. The Parkinson effect at the individual level appears to occur insidiously and without obvious benefit (Study 4) to work quality.

Although virtually all laboratory experiments carry with them implicit time norms, each of the three work sets in both of the present paradigms was easily accomplished in a few minutes by the present subjects. Parkinsonian work expansion may be generalizable to a wide variety of judgment, preference, classification, and monitoring tasks. Whenever anticipated work on the next task is canceled or, more generally, excess time arises, dalliance by workers on their present task could amount to substantial and costly inefficiency.

REFERENCES

- ARONSON, E., & GERARD, E. (1966). Beyond Parkinson's law: The effect of excess time on subsequent performance. *Journal of Personality & Social Psychology*, 3, 336-339.
- ARONSON, E., & LANDY, D. (1967). Further steps beyond Parkinson's law: A replication and extension of the excess time effect. *Journal of Experimental Social Psychology*, 3, 274-285.
- BRYAN, J. F., & LOCKE, E. A. (1967). Parkinson's law as a goal-setting phenomenon. *Organizational Behavior & Human Performance*, 2, 258-275.
- CARVELL, F. J. (1980). *Human relations in business*. New York: Macmillan.
- CONKLIN, J. E. (1984). *Sociology: An introduction*. New York: Macmillan.
- GIDDENS, A. (1989). *Sociology*. Cambridge, U.K.: Polity Press.
- HERSEY, P., & BLANCHARD, K. H. (1988). *Management of organizational behavior—Utilizing human resources*. Englewood Cliffs, NJ: Prentice-Hall.
- JOHNSON, B. T. (1993). *DSTAT 1.10: Software for the meta-analytic review of research literatures* (Version 1.10) [Computer software]. Hillsdale, NJ: Erlbaum.
- KOONTZ, H., & WEIRICH, H. (1990). *Essentials of management*. New York: McGraw-Hill.
- KORNBLUM, W. (1988). *Sociology in a changing world*. New York: Holt.
- LANDY, D., MCCUE, K., & ARONSON, E. (1969). Beyond Parkinson's law: III. The effect of protractive and contractive distractions on the wasting of time on subsequent tasks. *Journal of Applied Psychology*, 53, 236-239.
- LUTHANS, F. (1981). *Organizational behavior*. New York: McGraw-Hill.
- MACONIS, J. (1991). *Sociology*. Englewood Cliffs, NJ: Prentice-Hall.
- MCGRATH, J. E., & KELLY, J. R. (1986). *Time and human interaction: Toward a social psychology of time*. New York: Guilford.
- MINER, J. B. (1980). *Theories of organizational behavior*. Hinsdale, IL: Dryden.
- ORPEN, C., & RIESE, E. (1973). Beyond Parkinson's law: A failure to replicate. *Journal of Social Psychology*, 89, 151-152.
- PARKINSON, C. N. (1957). *Parkinson's law*. Cambridge, MA: Riverside Press.
- PARKINSON, C. N. (1980). *Parkinson: The law*. Boston: Houghton Mifflin.
- PETERS, L. H., O'CONNOR, E. J., POOYAN, A., & QUICK, J. C. (1984). The relationship between time pressure and performance: A field test of Parkinson's law. *Journal of Occupational Behaviour*, 5, 293-299.
- REIMANN, B. C. (1979). Parkinson revisited: A component analysis of the use of staff specialists in manufacturing organizations. *Human Relations*, 32, 625-641.
- ROBBINS, S. P. (1990). *Organization theory: Structure, design and applications*. Englewood Cliffs, NJ: Prentice-Hall.

THIO, A. (1991). *Sociology: A brief introduction*. New York: Harper-Collins.

NOTES

1. "In any public administrative department not actually at war, the staff increase may be expected to follow this formula: $x = (2km + 1)/n$, where k is the number of staff seeking promotion through the appointment of subordinates; 1 represents the difference between the ages of appointment and retirement; m is the number of manhours devoted to answering minutes within the department; and n is the number of effective units being administered, x will be the number of new staff required each year" (Parkinson, 1957, p. 12).

2. Time cuing affected other Parkinson studies as well. Bryan and Locke (1967) obtained an effect similar to Aronson's. Subjects who were given excess time took longer to complete a multitrial task than did subjects who were given minimum time. As in the work by Aronson, the instructions were replete with time cues and with the experi-

menter's announcements of temporal expectations. Orpen and Riese (1973) reported a conceptual replication that likewise included time cues. Peters, O'Connor, Pooyan, and Quick (1984) reported an ambiguous correlational study in which employees were asked about task times and the adequacy of those times.

3. In the absence of main and interaction effects of gender, a confounding of subject gender with the gender of people in photos—for example, during the test trial, male subjects saw photos of own-gender faces and females did not—could not account for the overall pattern of Parkinson effects.

4. "Dalliance" has the primary meaning of "a trifling away of time; dawdling" (unabridged *Random House Dictionary of the English Language*). In the present context, however, no pejorative connotation is attached to "dalliance" other than taking more time than necessary.

(Manuscript received July 7, 1997;
revision accepted for publication April 1, 1998.)