Age Differences in Transitory Cognitive Performance

Sy Miin Chow (smc9x@virginia.edu)
Department of Psychology; University of Virginia; P.O. Box 400400
Charlottesville, VA 22904-4400 USA

John R. Nesselroade (jrn8z@virginia.edu)
Department of Psychology; University of Virginia; P.O. Box 400400
Charlottesville, VA 22904-4400 USA

Abstract

Short-term performance data from a complex computerized cognitive test called SYNWORK1 were examined for age differences in transitory performance fluctuations in samples of 55 older and 57 younger adults. Profile analysis indicated that the older adults' performance trajectories were essentially parallel to those of the younger adults', but with the older adults performing at a consistently lower level on all four subtasks of SYNWORK1. These apparent age differences in level of performance were reduced substantially when a simple graphical approach was used to examine the performance trajectories. These results extend our knowledge concerning the nature of intraindividual variability while illustrating again some of the methodological inadequacies inherent in research comparing age differences in levels of cognitive performance when common statistical assumptions are even mildly violated. The competence of older adults can be underestimated based on a single measure of a group mean, thus leading to further risk of missing important learning strengths of older adults.

Selection and selection effects have received a considerable amount of attention from behavioral scientists (Nesselroade, 1988; Nesselroade & Thompson, 1995) and still remain one of the obstacles researchers must somehow overcome. The primary concerns, however, revolve around selecting a representative sample of participants from the population of interest (e.g. Cronbach, Gleser, Nanda & Rajaratnam, 1972) and valid indicators to represent the underlying construct under study (e.g. Little, Linderberger & Nesselroade, 1999). These, of course, capture only two of the ten possible dimensions defining empirical data in Cattell's data box (Cattell, 1966; Little et al., 1999), namely, the persons and variables dimensions, among other possible design configurations. Another relatively familiar dimension of the data box, occasions of measurement, has also been discussed rather extensively, especially in comparing the relative merits of cross–sectional versus longitudinal research design (e.g. Kraemer, Yesavage, Taylor & Kupfer, 2000). Another kind of selection effect that is inherent in almost any research designs, but has rarely been addressed, is the effect of averaging data across participants or occasions of measurement. In a recently published article, Newell, Liu and Mayer–Kress (2001) question the common practice of averaging data across participants or occasions, presumably to remove the transient, noise–like changes from trial–to–trial, or during the "warm–up" phase at the beginning of a practice session, with the goal of singling out a global learning trend that is characteristic of all the participants across all the trials. As suggested by Lamiell (1981), both idiographic and nomothetic approaches have their own merits in answering certain research questions. However, when a group mean is used as the only index of a group's performance, the end of searching for a global trend in learning does not always justify the means of levelling out the individual differences in this aspect.

Idiographic and Nomethetic Approaches to Modeling Change

Over the past few decades, the importance of an idiographic approach (Allport, 1937; Murray, 1938) to studying human behavior has gained increased recognition. Considerable efforts have been devoted to integrate idiographic and nomothetic approaches in psychological research, thus allowing researchers to capture both the intraindividual variability, and the interindividual differences in various aspects of human behavior (Baltes & Nesselroade, 1979). Repeated assessments of the same individual often yield information on intraindividual variability in aspects thought to be relatively stable over short time-span, such as cognitive abilities and intelligence (see e.g., Horn, 1972; May, Hasher & Stoltzfus, 1993; Stigler, 1994), personality styles and other belief systems (e.g., Shoda, Mischel & Wright, 1994; Kim, Nesselroade & Featherman, 1996), as well as other more transient state-like fluctuations in affective states (e.g., Larsen, 1987; Shifren, Hooker, Wood & Nesselroade, 1997; Mumma 2001).

While many researchers are moving away from performing means comparisons at the aggregate level, the idea of taking a group mean as the unbiased estimator of the group, as well as the population that it represents, is so deeply entrenched in contemporary data analytic techniques that a majority of the between–group comparisons essentially involve means. Although it is a well–known fact that other measures of central tendency, such as the median, might be a better estimator of a group's performance when certain statistical criteria are not met, or when there are outliers in the data that could potentially skew the results of one's analysis, most of the available statistical tests, such as ANOVA, are based on means comparisons. Even in cases where intraindividual variability is modeled explicitly, such as in growth curve analysis (McArdle & Epstein, 1987; Francis, Fletcher, Stuebing, Davidson & Thompson, 1991; Meridith & Tisak, 1990), conclusions on intraindividual changes in levels of performance still derive primarily from means. Adding to these methodological difficulties of capturing representative idiographic information, of course, are the conceptual difficulties of summarizing the interindividual differences in a way that is helpful for making empirical decisions. A graphical approach is a useful supplement to other more rigorous statistical approaches, as it helps to depict a summary picture of both the intraindividual and interindividual aspects of change.

Transitory Changes in Cognitive Performance as Meaningful Intraindividual Changes

Despite increased awareness of the limitations of using an aggregate measure to represent a group, the same limitations that exist when applied to individual data were not addressed as often. Just as a group mean does not necessarily represent the group as a whole, an individual's mean score is limited in its own way. The transient fluctuations observed during the initial phases of an individual's learning history should not unthinkingly be regarded as "outliers" that ought to be levelled out. Unfortunately, most experimental studies aimed at capturing deterministic dynamics in transitory learning fluctuations as observed during the "warm-up" phase are limited to studies in the area of motor development (e.g. Adams, 1961; Schmidt, 1982; Thelen, 1994). Individuals' cognitive learning curves as reported by some researchers (e.g. Salthouse, Hambrick, Lukas & Dell, 1996) do in general resemble the learning curve of motor skills. However, very few studies have focused on examining the patterns of transitory changes in human cognitive performance and even fewer studies emphasize learning as an ongoing refinement of errors in the face of external perturbations.

Objectives of This Study

In this paper, we examine transitory changes in individuals' short-term adaptive responses and compare these responses among adults of different ages. We also demonstrate some of the inadequacies inherent in comparing adults of different ages on the basis of their group means when common statistical assumptions are even mildly violated. Finally, we present a simple graphical approach that serves as a supplement to means comparisons at a group level, and an alternative to summarizing changes at an individual level.

Method

Participants

This sample consisted of 55 older adults (36 female and 19 male), and 57 younger adults (38 female and 19 male). The older adults' ages ranged from 60 to 93 years ($M = 73.73$, $SD = 7.26$), and the younger adults' from 17 to 28 years ($M = 19.18$, $SD = 1.75$). The younger adults were recruited from the undergraduate participant pool in a southeastern university and were rewarded with course credits for participation in this study. The older adults were recruited from the Charlottesville community through newspaper advertisements and direct solicitation at senior centers . Both age groups rated themselves as in reasonably good health, with self-ratings of health averaged between average to good (on a 4-point scale from 1—excellent to 4—poor, $M = 2.27$ and $SD = 1.00$ for older adults; $M = 1.91$ and $SD = 0.87$ for younger adults).

Materials

The cognitive performance of the participants was assessed using a computer program called SYNWORK1. The SYNWORK1 program is a computerized multi–tasking test environment designed by Elsmore (1994) to examine an individual's ability to perform multiple tasks concurrently. Figure 1 shows the four primary tasks that are measured on SYNWORK1, including a memory task, a self–paced arithmetic task that involves using mouse–click to manipulate the plus and minus panels to adjust the sum from four zeros (0000) into the correct sum of two given numbers, a visual monitoring task, and an auditory discrimination task. Points are given for correct responses and taken away when a task is neglected or performed incorrectly. An individual's total score is constantly updated and is shown in the middle of the four task quadrants.

Procedure

Due to the complexity of the SYNWORK1 program, the participants were first given a short training session during which they practiced the four SYNWORK1 tasks one at a time, followed by a one–minute session during which they practiced the four tasks simultaneously. After the participants were told to strive for their best possible scores, we began recording their performance data on SYNWORK1 over nine consecutive trials, with each trial lasting about one minute.

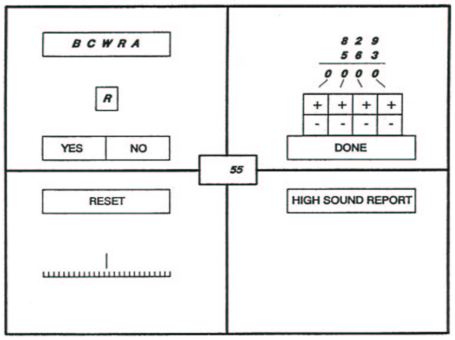

Figure 1: The SYNWORK1 program designed by Elsmore (1994) to test human synthetic work performance

To induce some short–term transitory fluctuations in the participants' performance, we manipulated the difficulty levels of the SYNWORK1 program throughout the nine trials by speeding up or slowing down the timing at which each task appeared. During the more difficult sessions, we also increased the number of addends for the arithmetic task, and the number of target letters the participants had to remember for the memory task. This sudden increase in the difficulty levels of the four tasks was expected to induce a small amount of perturbations in the individuals' responses and the individuals' ability to adapt to these sudden perturbations was taken as an indication of their adaptive behavioral patterns. During a particular trial, the SYNWORK1 program was governed by one of three possible sets of parameters we classified as easy, mid–difficulty, and difficult. The participants were exposed to one of two sequences of altering difficulty levels in the order of {M, E, M, D, M, E, M, D, M}, or {M, D, M, E, M, D, M, E, M}, where M represents mid–difficulty, E represents easy, and D represents difficult sessions. The total number of correct memory, arithmetic, visual, and auditory responses for each of the nine trials were used as the primary dependent variables in this study to determine if the two age groups had different performance profiles on each of the tasks.

Results

Prior to analysis, the participants' task scores over nine trials were screened for outliers and departure from other statistical assumptions. Several outliers were detected among younger adults who performed too poorly on the memory task, older adults who performed too well on the arithmetic task and the visual task, and among younger adults on the auditory task due to the restricted range of scores observed in that age group. These outliers were retained in subsequent data analysis as they were thought to reflect reasonable range of fluctuations in performance. No missing values were observed in the data. When subject to a MANOVA test, the task scores of individuals exposed to the two sequences of difficulty levels were not statistically different from each other. After confirming that, a profile analysis was used to evaluate the differences in level, parallelism, and flatness of the two groups' responses on the four tasks, with age group as the independent variable.

Profile Analysis

The descriptive statistics of the two age groups are presented in Table 1. Consistent with findings reported in the aging literature, the younger adults were found to perform at a significantly higher level on the memory task, $F(1{,}110) = 132.17$, $p < .001$; the arithmetic task, $F(1{,}110) = 223.69$, $p < .001$, and the auditory task, $F(1{,}110) = 106.56$, $p < .001$, as determined by using Wilk's criterion. The visual task was the only task that did not show significant difference in levels between the two age groups.

Table 1: Descriptive Statistics of the Older and Younger Adults.

	Mean		**Standard Deviation**	
	Old	Young	Old	Young
Memory	5.77	9.47	2.92	2.20
Arithmetic	0.21	1.84	0.60	1.45
Visual	70.80	72.22	22.51	21.49
Auditory	0.81	1.78	0.83	0.76

When averaged across groups, scores on all four tasks were found to deviate significantly from flatness by Hotelling's criterion on all four SYNWORK1 tasks ($p < .01$ for all). This simply confirmed that the experimentally imposed perturbations did lead to some fluctuations in the participants' performance. There might also be some learning taking place, as shown by the slight increase in means over the nine trials. Using Wilk's criterion, however, the older and younger adults' profiles were not found to deviate significantly from parallelism, except on the arithmetic task, $F(8,\ 103) = 2.98$, $p = .005$. The arithmetic task might be too cognitively demanding for the older adults, thus causing them to avoid the task altogether. Aside from this, the two age groups did not react to the experimental perturbations in a statistically different way over the nine trials.

Graphical Representation of Intraindividual Variability

The conceptual difficulties of summarizing information at an individual level can be illustrated using Figure 2, in which all the individuals' total scores on

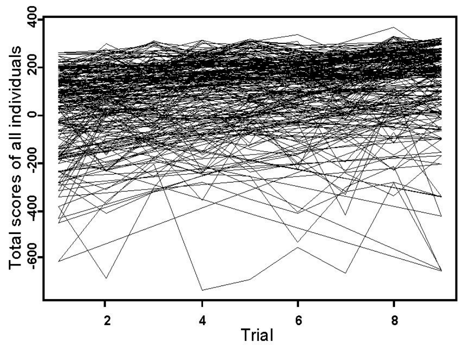

Figure 2: The total scores of all the individuals in the sample.

SYNWORK1 are shown over the nine trials. Different individuals have their own distinct trajectories, making it hard to summarize the performance of the group as a whole. The conventional means plot (see Figure 3) may be helpful in providing a summary picture, but could also mask important information in the data. To further examine the performance of the two age groups as captured by other descriptive measures, the 10th, 50th and 90th percentiles of the two groups' scores on the four tasks were plotted and compared using S–Plus' Hmisc library (Harrell & Alzola, 2001). Due to space limitations, we chose to include only the means and percentiles plots for the memory and auditory tasks here.

As shown in the plots in Figure 4, the age differences in performance levels reduced substantially when the percentiles plots were used to represent the performance of the two age groups [1]. When compared to adults from the same age group, older adults who performed at the 90th percentiles on the memory and arithmetic tasks (omitted here) were found to show performance that closely resembled younger adults whose scores were near the median levels of their peers[2]. Older adults in this category were observed to make more marked improvements on the memory and arithmetic tasks toward the end of the experiment. A similar pattern was observed on the auditory task, except that the relatively well performing older adults (i.e., the 90th percentile group), started out with performance trajectory that was identical to the trajectory of younger adults who performed at the median level, but after the 7th trial, caught up to the performance of younger adults who performed at the 90th percentile level and performed at exactly the same optimal level after the 7th trial. The younger adults who performed

[1] We also plotted the trajectories of younger and older adults at less extreme percentile levels (e.g. 25th and 75th percentiles). A similar, but smaller magnitude of reduction in age differences was observed.

[2] The percentile levels were calculated separately for each task and each age group.

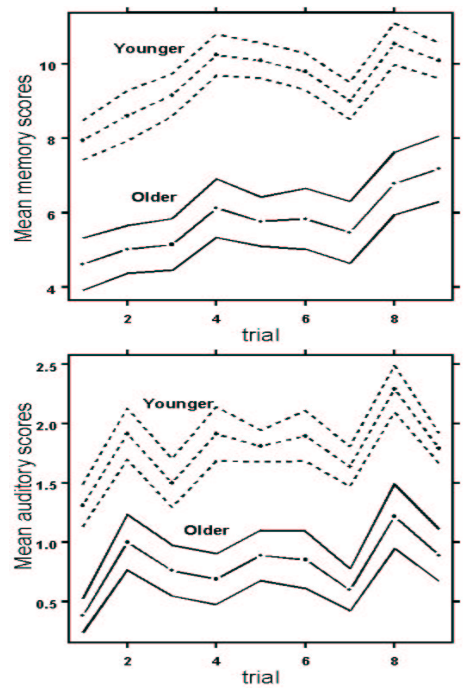

Figure 3: The means and the 95% confidence limits of the older and younger adults' scores on the memory and auditory tasks.

at the 90th percentile levels were performing at the optimal level on the memory, visual and auditory tasks very early on, achieving the maximum possible scores during almost all of the trials. Their only source of improvement in scores derived primarily from the self–paced arithmetic task, on which their scores continued to improve throughout the course of this experiment. On the other hand, older adults at the 90th percentile level were observed to have rather uniform improvements on all four tasks, and more trials were required before they could attain the same level of performance the well performing younger adults could achieve at an early phase.

Another important source of variability constituted by age stemmed from the lack of clear improvements observed among older adults who performed at the 10th percentile level. The relatively poor performing younger adults (those at the 10th percentile level) were able to capitalize on the memory task and auditory task as time progressed much better than older adults of comparable performance level. Nevertheless, older adults at the 10th percentile level did demonstrate their own learning strengths on the visual task—they attained a considerable amount of improvement on the visual task from trial one to trial two, and maintained a rather persistent level of performance before a sharp decrement in scores was observed during and after the 8th trial, presumably due to fatigue and decreased attention span. When the

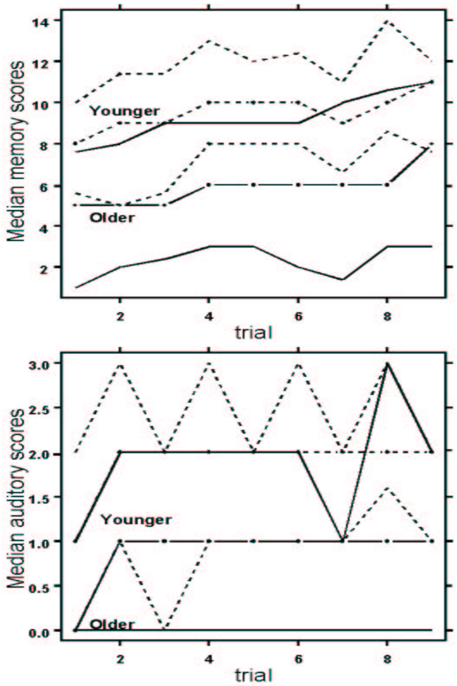

Figure 4: The 10th, 50th and 90th deciles of the older and younger adults' memory and auditory scores.

plots of individuals from less extreme percentile levels (e.g. 25th and 75th percentiles) were examined, similar fluctuation patterns that differed slightly in levels were observed, revealing interesting individual differences in performance fluctuations within, as well as between different age groups.

Discussion

In this study, SYNWORK1 was used as an active interface for capturing the age differences in performance fluctuations when individuals were faced with ongoing external perturbations. As Jones and Conrad's (1933) quotation of Thorndike's remark put it, "...individual differences amongst those of the same age...enormously outweigh differences between ages" (p. 258–59). The common approach of comparing age differences in levels of cognitive performance by using group means inevitably under–represents the complexity underlying the variability in performance both within and between different age groups, especially when the two age groups have unequal variances in many respects. Due to the lack of a clear definition of what constitutes a normative representation of the older adults population, the issue of identifying and eliminating outliers becomes tricky. Results from our profile analysis showed findings that were consistent with those reported in the literature (e.g. Erber, 1976). Essentially, adults of different age groups were found to exhibit similar, significant improvements on cognitive or intelligence tests as practice effects accumulated, but the younger adults' level of performance was almost always higher than the older adults' on all the measurement occasions.

Using group means and the changes in means as the sole indicators of the performance of these older and younger adults might be informative in its own way, as demonstrated by the profile analysis in this study, but fails to acknowledge the differences in performance profiles among individuals of the same age group. In fact, the younger adults in this study, who were all recruited from the same university and were often thought of as representing a rather homogeneous group, showed different dynamics in their performance fluctuations when the trajectories of individuals from different percentiles were compared. The trajectories plotted using the percentiles of the younger and older adults also revealed very consistent patterns in performance fluctuations that reflected our experimentally imposed alterations in task difficulty levels very accurately. More importantly, the age differences in levels of performance were reduced substantially when the percentile scores of these two age groups were examined, revealing some of the older adults' unique strengths in learning that started surfacing at a relatively more gradual pace than for the younger adults.

Of course, younger adults of the higher ability group might be encountering ceiling effects on those tasks from a very early phase. In addition, with the data from the present study, there is no way for us to determine whether the high–performance younger adults will resemble the high–performance older adults in any way when they get older. However, by using a simple graphical approach, we presented some of the inadequacies of using group means as the only representation of the dynamics of the group as a whole because a researcher may risk bypassing some of the interesting dynamics within the group by not looking at the results offered by other alternative methods.

Acknowledgments

This research was supported by the National Institute Aging grant 5 R01 AG18330 awarded to the second author.

References

Adams, J. A. (1961). The second facet of forgetting: A review of warm–up decrement. *Psychological Bulletin*, *58*, 257–273.

Allport, G. W. (1937). *Personality: A psychological interpretation*. New York: Holt.

Baltes, P. B., & Nesselroade, J. R. (1979). History and rationale of longitudinal research. In J. R. Nesselroade & P. B. Baltes (Eds.), Longitudinal research in the study of behavior and development. New York: Academic Press.

Cattell, R. B. (1966). The data box: Its ordering of total resources in terms of possible relational systems. In R. B. Cattell (Ed.), *Handbook of Multivariate Experimental Psychology* (1st ed.). Chicago, IL: Rand McNally.

Cronbach, L. J., Gleser, G. C., Nanda, H., & Rajaratnam, N. (1972). *The dependability of behavioral measures.* New York: Wiley.

Elsmore, T. F. (1994). SYNWORK1: A PC–based tool for assessment of performance in a simulated work environment. *Behavior Research Methods, Instruments, & Computers*, *26*(4), 421–426.

Erber, J. T. (1976). Age differences in learning and memory on a digit–symbol substitution task. *Experimental Aging Research*, *2*(1), 45–53.

Francis, D. J., Fletcher, J. M., Stuebing, K. K., Davidson, K. C. & Thompson, N. M.(1991). Analysis of change: Modeling individual growth. *Journal of Consulting and Clinical Psychology*, *59*(1), 27–37.

Harrell, F. E. & Alzola, C. F. (2001). *An introduction to S–Plus and to the Hmisc design libraries.* Unpublished manual.

Horn, J. L. (1972). State, trait and change dimensions of intelligence. *British Journal of Educational Psychology*, *42*, 159–185.

Jones, H.E., & Conrad, H.S. (1933). The growth and decline of intelligence: A study of a homogeneous group between the ages of ten and sixty, *Genetic Psychology Monographs*, *13*.

Kim, J.E., Nesselroade, J. R. & Featherman, D. L. (1996). The state component in self-reported worldviews and religious beliefs of older adults: The MacArthur successful aging studies. *Psychology and Aging*, *11*(3), 396–407.

Kraemer, H. C., Yesavage, J. A., Taylor, J. L. & Kupfer, D. (2000). How can we learn about developmental processes from cross–sectional studies, or can we? *American Journal of Psychiatry*, *157*, 163–171.

Lamiell, J. T. (1981). Toward an idiothetic psychology of personality. *American Psychologist*, *36*(3), 276–289.

Larsen, R. L. (1987). The stability of mood variability: A spectral analytic approach to daily mood assessments. *Journal of Personality and Social Psychology*, *52*(6), 1195–1204.

Little, T. D., Lindenberger, U. & Nesselroade, J. R. (1999). On selecting indicators for multivariate measurement and modeling with latent variables: When "good" indicators are bad and "bad" indicators are good. *Psychological Methods*, *4*(2), 192–211.

May, C. P., Hasher, L. & Stoltzfus, E. R. (1993). Optimal time of day and the magnitude of age differences in memory. *Psychological Science*, *4*(5), 326–330.

McArdle, J. J. & Epstein, D. B. (1987). Latent growth curves within developmental structural equation. *Child Development*, *58*(1), 110–133.

Meridith & Tisak (1990). Latent curve analysis. *Psychometrika*, *55*, 107–122.

Mumma, G. H. (2001). Increasing accuracy in clinical decision making: Toward an integration of nomothetic-aggregate and intraindividual-idiographic approaches, *The Behavior Therapist*, *24*(4), 77–94.

Murray, H. A. (1938). *Explorations in personality.* New York: Oxford University Press.

Nesselroade, J. R. (1988). Some implications of the trait–state distinction for the study of development over the life span: The case of personality. In P. B. Baltes, D. L. Featherman & R. M. Lerner (Eds.), *Life–span development and behavior* (Vol 8). Hillsdale, NJ: Lawrence Erlbaum Associates.

Nesselroade, J. R. & Thompson, W. W. (1995). Selection and related threats to group comparisons: An example comparing factorial structures of higher and lower ability groups of adult twins. *Psychological Bulletin*, *117*(2), 271–284.

Newell, K. M., Liu, Y,& Mayer–Kress, G. (2001). Time scales in motor learning and development. *Psychological Review*, *108*(1), 57–82.

Salthouse, T. A., Hambrick, D. Z., Lukas, K. E. & Dell, T. C. (1996). Determinants of adult age differences on synthetic work performance. *Journal of Experimental Psychology: Applied*, *2*(4), 305–329.

Schmidt, R. A. (1982). *Motor control and learning.* Champaign, IL: human Kinetics.

Shifren, K., Hooker, K., Wood, P. & Nesselroade, J. R. (1997). Structure and variation of mood in individuals with Parkinson's disease: A dynamic factor analysis. *Psychology and Aging*, *12*(2), 328–339.

Shoda, Y., Mischel, W. & Wright, J. C. (1994). Intraindividual in the organization and patterning of behavior: Incorporating psychological situations into the idiographic analysis of personality. *Journal of Personality and Society Psychology*, *67*(4), 674–687.

Siegler, R.S. (1994). Cognitive variability: A key to understanding cognitive development. *Current Directions in Psychological Science*, *3*, 1–5.

Thelen, E. & Smith, L. B. (1994). *A dynamic systems approach to the development of cognition and acion.* Cambridge, MA: MIT Press.