The Structure of Individual Differences in Batteries of Rapid Acquisition Tasks in Mice

Charles Locurto, Andrea Benoit, Caitlin Crowley, and Andrea Miele
College of the Holy Cross

Two experiments examined the structure of individual differences in mice by means of tasks that produced significant acquisition within 1 session. In Experiment 1, 5 cognitive tasks—detour, winshift, olfactory discrimination, fear conditioning, and operant acquisition—were used in conjunction with two control procedures: an open field and a light–dark test. In Experiment 2, some modifications were made to the tasks used in the 1st experiment, and 3 new tasks were used in conjunction with the same control procedures. The battery consisted of 5 learning tasks: detour, Hebb–Williams, radial maze, olfactory foraging, and fear conditioning. Results of both experiments indicate that when cognitive tasks and control procedures were included in principal-components analyses most of the variance attached principally to individual tasks rather than to a general component as is found typically in human cognitive batteries. When control procedures were eliminated, there was better evidence for the presence of a general cognitive factor, particularly in Experiment 2.

Keywords: individual differences, cognition, acquisition, mice

When human subjects are given a battery of cognitive tasks, they tend to retain their rank ordering across those tasks, thereby producing a matrix of positive correlations. When factored, this matrix yields a first factor that accounts for an appreciable proportion of matrix variance, ranging from one third to one half or more of total variance (Jensen, 1998; Locurto, 1997; Miller & Vernon, 1992). This finding has encouraged the view that in addition to the specific processes evoked by particular tasks there may be a general process or set of related processes that are involved in the solution of a wide range of cognitive tasks. To the extent that a general process can be identified, it stands in contrast to the view currently found among evolutionary psychologists that cognitive abilities are best characterized as modular, having arisen in response to specific evolutionary demands (e.g., Barlow, Cosmides, & Tooby, 1992; Tooby & Cosmides, 2000).

The nonhuman animal literature is meager compared with the human literature on this issue. This neglect derives perhaps from the prominence of ecological approaches to animal behavior. From the perspective of these approaches, much like the assumptions underlying evolutionary psychology, the evolution of cognitive mechanisms is assumed to have been driven by problems confronted within the particular natural history of a species (see Collias & Collias, 2004; Gallistel, 2003; Seligman & Hager, 1972). If true, selection pressures would tend to encourage the emergence of niche-specific mechanisms, termed *adaptive specializations*, rather than general process mechanisms. An alternative view, developed by Macphail (1987), argues that cognitive mechanisms may be uniform across species and, perhaps, individuals. Although this position differs from an adaptive specialization approach, it also does not expect the presence of systematic individual differences across different tasks.

Those studies that have addressed individual differences in animals do not at present yield a uniform conclusion. Anderson (1993) observed positive correlations from a battery of three cognitive tasks in rats and suggested that this result was evidence of a general process mechanism. Locurto and Scanlon (1998) reported a similar finding across a series of water escape tasks in mice. These two studies, although demonstrating that a matrix of positive correlations can be observed in a nonhuman species, were each limited by the use of the same motivational system across all tasks. The positive correlations might therefore be due to motivational rather than to cognitive commonalities.

More recent work that expanded the types of tasks studied has yielded a mixed picture. Locurto, Fortin, and Sullivan (2003) developed a battery of six tasks that used both appetitive and aversive motivational systems. Results indicated a modular component structure in which the six tasks produced three components each accounting for 16%–20% of variance. Conversely, in two other studies in which mice were used, researchers have observed matrices of more uniformly positive correlations. Galsworthy, Paya-Cano, Monleon, and Plomin (2002) used a battery of six tasks that spanned appetitive and aversive motivational systems. They observed positive correlations and a first component that accounted for approximately 30% of variance. (See Galsworthy et al., 2005, for a similar though somewhat weaker finding with respect to a general process.) A recent study by Matzel et al. (2003) reported a similar finding.

One difference between the study by Locurto et al. (2003) and those of Galsworthy et al. (2002) and Matzel et al. (2003) concerns the manner in which the individual tasks were implemented.

Locurto et al. ran each task for a number of sessions until responding reached a criterion that was at or near asymptote. Galsworthy et al. used a mixture of training durations, some of which were only one or a few sessions. Matzel et al. exclusively used tasks that were run for one or a few sessions. In other words, the batteries in Galsworthy et al.'s and Matzel et al.'s studies tapped principally acquisition, whereas those in Locurto et al.'s study measured behavior at or near an asymptotic level of performance.

It is unclear from the literature on acquisition and performance whether these two aspects of training should be regarded as similar. Some findings suggest that *acquisition*, defined as responding early in training, and *performance*, defined as responding following acquisition, are experimentally dissociable. Other findings suggest that these two aspects of training respond similarly to the same experimental interventions (cf. Landeira-Fernandez, Fanselow, DeCola, & Kim, 1993, and Verrico, Jentsch, Roth, & Taylor, 2004, with Jakubowska, Gumusbas, & Kara, 2003, and Nikkhah, Rosenthal, Hedrich, & Samii, 1998). These studies all measured behavior at two different points in training to determine the relationship between acquisition and performance. As a consequence, they can provide only indirect evidence concerning whether individual differences in a battery composed solely of tasks run for a brief period of acquisition would produce results similar to those found in a battery in which tasks were run for longer durations.

In the two experiments reported herein, batteries were developed that were composed solely of tasks that evidenced acquisition within one session. The question of interest was whether batteries constructed in this manner would yield clearer evidence of a general factor as compared with the results of Locurto et al. (2003).

Experiment 1

Method

Subjects

Forty-seven heterogeneous stock mice were obtained from the Institute for Behavior Genetics, University of Colorado. These mice are an outbred strain that was originally derived from an eight-way cross of inbred mouse strains (see McClearn, Wilson, & Meredith, 1970). Subjects were approximately 90 days old at the start of training and were approximately 165 days old at the end of training. They were housed in individual polycarbonate cages (Ancare, Bellmore, NY) that measured 29.5 × 15 × 10.5 cm. They were maintained on a 12-hr reversed day–night cycle with food and water freely and continuously available prior to the start of training and during water escape and fear-conditioning tasks. They were tested at least 1 hr into their dark cycle. Subjects were run in two samples. The first sample consisted of 20 subjects; the second sample consisted of 35 subjects. Two subjects died during testing. The data from 6 additional subjects were not used in any analyses because they evidenced satiation effects during the olfactory task and did not complete the six trials of that task. The data from the remaining 47 subjects were used in all analyses.

Apparatus, Task Descriptions, and Procedures

The battery consisted of two control procedures and five cognitive tasks. The control procedures were used to assess subjects' preexperimental behaviors with respect to activity, exploration, and anxiety. These procedures consisted of an open field test and a light–dark test. The cognitive tasks consisted of olfactory discrimination, a winshift problem, a detour problem, a nose-poke acquisition procedure, and fear conditioning. Each task was designed to tap a different response system, that is, a different set of response type, sensory modality, and motivation. Subjects experienced these tasks in the order here presented with the exception that the nose-poke acquisition procedure and fear conditioning were administered in counterbalanced fashion between the two samples, with fear conditioning administered as the last task for the first sample, and operant acquisition administered as the last task for the second sample.

Open field. This task was conducted in a square arena constructed of Plexiglas that measured 61 cm on each side and was 46 cm high. Subjects were given one 10-min session. Lighting consisted of one 90-watt fluorescent tube lamp positioned approximately 2.5 m above the arena and recessed into the ceiling (122 lux as measured from the center of the arena). Movements were recorded via a video-monitoring system (EthoVision, Noldus, Wageningen, The Netherlands). Four measures were taken from this test: total distance covered, amount of time spent hugging the walls (thigmotaxis), center crossings, and rearings.

Light–dark test. This procedure was run in an arena measuring 61 cm square and 37 cm high (TruScan, Coulbourn Instruments, Allentown, PA). The arena was divided into light and dark halves. The dark half consisted of a four-sided insert made of black Plexiglas that covered one half of the arena and the wall to the rear of the insert. The light side was illuminated with a 90-watt fluorescent lamp positioned approximately 2.5 m above the arena (284 lux). The task was run for one 10-min session. Two measures were taken from this test: the proportion of time spent on the light side of the chamber and the number of transitions between the light and dark areas of the chamber.

Olfactory discrimination. This task was adapted from Mihalick, Langlois, Krienke, and Dube (2000). Mice were tested in a polycarbonate cage that measured 56 × 25 × 22 cm. A T-shaped soda cap holder was situated at one end of the cage. The cap holder was made of black Plexiglas with two inset holes 3.5 cm in diameter in its base that were separated by a vertical rectangular barrier 12 cm in length and 10.5 cm high. Each hole held a plastic soda bottle cap. The caps were filled with corncob laboratory bedding that was scented by mixing .6 g of either cinnamon or nutmeg (McCormick, Sparks, MD) with 100 g of bedding. A .2-g piece of semisweet chocolate (Nestlé, Glendale, CA) served as the reinforcer and was buried beneath the top of the bedding in one of the scented caps during training.

On each trial the subject was started from the opposite end of the cage from the scented bedding caps. Each trial began with a mouse placed in a cylinder 16.5 cm high and 10.5 cm in diameter. The cylinder had a small curved escape tunnel 5.5 cm in length at its base that funneled the mouse into the center of the cage. The tunnel could be blocked with a piece of clear Plexiglas until the trial began. This system was used to orient the mouse toward the two scented stimuli at the beginning of each trial. The barrier between the two scents was used to prevent quick crossovers from one cap to the other.

Training began by depriving the mice of food for 24 hr. On the following day they received 5 g of standard laboratory chow, followed by 4 g the next day and by 3.5 g daily thereafter. This procedure reduced the subjects to 80%–90% of their free-feeding weight. Mice were also given .2-g pieces of chocolate buried in unscented bedding in a soda cap in their home cages on the 3rd day. On the 4th day they were habituated to the test chamber by allowing them to explore the testing arena. One cap filled with plain bedding and with three pieces of chocolate—one above the bedding and the other two slightly buried in the bedding—was placed in one of the soda caps in the T-shaped cap holder. After the subject consumed the three pieces in that one cap, three additional trials were run in which the placement of the cap, now baited with one piece of chocolate, was alternated between the left and right caps.

On the following day olfactory discrimination was conducted. Each subject received six trials for one session. Cinnamon was associated with

the reinforcer, and its location was randomly alternated between the left and right caps, with the provision that the correct scent occurred in a given position on no more than three consecutive trials. An error occurred when the bedding associated with the nonreinforced scent was disturbed. Sniffing was not counted as an incorrect response. Each subject had 40 s to locate the reinforcer on each trial. As noted earlier, 6 subjects did not complete the six trials of this task and were eliminated from analysis.

For all tasks, subjects were run in subgroups of 3 or 4, such that each subject experienced the first trial before the second trial was administered to any subject. This procedure had the effect of ensuring that the average intertrial interval (ITI) ranged between 60 and 180 s. Subjects were placed in holding cages between trials. These were standard mouse cages with bedding on the floor but without food or water. Illumination, provided by one overhead fluorescent light, measured 310 lux from the bottom of the testing arena. The arena was cleaned between trials with a 50% water/50% ethanol solution. Performance was measured as the average number of errors recorded over the first four trials. This criterion was adopted because several subjects evidenced large increases in latency and errors over the final two trials, suggesting that satiation had taken place. As earlier noted, 6 subjects failed to complete these final two trials.

As this procedure places the chocolate reinforcer in with the correct scent, there is the possibility of a confound in the form of subjects being able to detect the buried chocolate irrespective of the scent. In several tests of this confound we have found that when the chocolate is paired randomly with each scent, performance does not rise above chance. (See, also, Mihalick et al., 2000, for further evidence on this point.)

Winshift. This task was run under water escape motivation. The procedure was conducted in a Y maze. Each arm measured $46 \times 31 \times 15$ cm. One arm of the maze served as the start arm for all trials. The remaining two arms served as choice arms. A metal platform $15 \times 10 \times 16.5$ cm submerged just below the water line served as the escape platform. The maze was set in a children's wading pool. Water temperature was maintained at 18 °C ± 1 °C. A single session of habituation preceded conditioning. In the habituation session each subject received four forced-choice trials. In these trials one of the choice arms was blocked, and the subject was forced to choose the one open arm. The escape platform was located at the end of the open arm. Four trials were conducted in this manner for each subject. On finding the escape platform, the subject remained on the platform for 10 s. If the subject did not find the platform within 40 s, it was placed on the platform for 10 s. The procedure of ending a trial with the subject on the platform was followed during conditioning. Trials were run such that each subject experienced its first trial before the first subject experienced the second trial.

During conditioning, a trial consisted of two runs. The first run was a forced-choice run in which only one choice arm was open and contained an escape platform. After the subject completed that run, it was placed in a holding cage that had paper strips on the bottom that were replaced as needed during a session. After approximately a 10-s delay, the subject received a choice trial in which both arms were open and the escape platform was placed in the formerly blocked arm. The platform during free-choice trials was placed equally often in the left or right arm relative to the start arm. Subjects received eight trials in one session. Errors consisted of a subject passing the midpoint of the incorrect arm or returning past the midpoint of the start arm. Indirect lighting in the test room was provided by three 75-watt lamps (53 lux, measured from the center of the arena). Following each trial, the water was whisked to remove debris and to disrupt odor trails.

Detour. This problem was run in a rectangular wooden arena that was $115 \times 39 \times 38$ cm. One end of the arena was painted black and served as the goal area. The other end of the arena was painted off-white and served as the start area. The top of the goal area was covered by a black cloth. Various toys, including tunnels and a running wheel, were placed inside the goal area. When subjects reached the goal area they were given 3 min of exploration. One session of habituation preceded conditioning. During

habituation, subjects were given three trials. On the first trial, the arena contained no barriers, and the subject was allowed to traverse the arena to reach the goal area. On the second trial, a floor-to-ceiling barrier was placed in front of the start area to confine the subject before the trial began. There was a similar barrier in front of the goal area with a 7.6-cm square hole cut at floor level that constituted the only entry into the goal area. On the third habituation trial, these two barriers were used in conjunction with an inclined plane. The inclined plane was an insert 8 cm at its high end and floor level at its low end, 38 cm wide, and 35 cm long. The plane was placed in the center of the area sideways to allow the subject to crawl over it to reach the goal area.

During conditioning trials on the next day, the plane was turned so that its high end faced subjects as they left the start area. Subjects had to mount the plane, then descend along it into the square opening at floor level to reach the goal area. Any movement that resulted in the subject's body except its tail crossing and then recrossing an imaginary line formed by the leading edge of the detour constituted an error. Errors were scored by observers from a video monitor. Lighting consisted of one 90-watt fluorescent tube lamp positioned approximately 2.5 m above the arena and recessed into the ceiling (122 lux as measured from the center of the arena). The detour arena was cleaned between trials with a 50% water/50% ethanol solution.

Nose-poke acquisition. This task was conducted in operant conditioning chambers. The subject's portion of the chamber was $25.4 \times 14.6 \times 28$ cm. A nose-poke that was illuminated with green light and a feeder that delivered .02 ml of water were housed on the front panel (Coulbourn Instruments, Allentown, PA). One session of feeder training preceded the conditioning session. During feeder training subjects were required to earn 10 reinforcements within a 60-min period. Food was delivered following an entry into the feeder. On the day preceding feeder training, subjects had water removed from their home cages. Following feeder training, subjects received 30 min of access to water in their home cages. Acquisition of the nose-poke response took place on the following day in one 90-min session. The acquisition criterion required the subjects to earn 25 reinforcements that were delivered on a fixed-ratio one schedule.

Fear conditioning. This procedure was conducted in a chamber that measured $25 \times 20 \times 28$ cm. The conditioned stimulus (CS) was a 1000-Hz tone presented for 30 s. Sound pressure at the center of the chamber measured 93 dB during the tone, against an ambient background noise level provided by a vent fan of 76 dB. The unconditioned stimulus (US) was a .5 mA footshock presented for .5 s. Conditioning was conducted in a single session. The session began with a 5-min habituation period. Thereafter four CS–US presentations occurred. The ITI was 3 min. Activity during the tones was measured by an infrared activity monitor (Coulbourn Instruments, Allentown, PA) placed in the ceiling of the chamber. The dependent measure for this task was the proportionate reduction in activity between first and last trial [(first trial activity − last trial activity)/first trial activity)]. This measure was then scaled in the same direction as errors and operant acquisition times by multiplying scores by −1 so that larger proportionate reductions in activity produced smaller ratios.

Results and Discussion

The control tasks were analyzed separately to reduce the six dependent measures derived from the two tasks into a smaller subset of components. Principal-components analysis (SPSS 11.0, no rotation) of these measures yielded two components with eigenvalues greater than 1.0. The first component accounted for 30% of variance and was marked principally by the distance measure from the open field (loading = .81) and the number of transitions in the light–dark test (loading = .73). This factor was termed *activity*. The second factor also accounted for 30% of matrix variance and was marked most strongly by the difference between

the thigmotaxis measure taken from the open field (loading = .86) and the number of center crossings in the open field (loading = −.76). This component was termed *anxiety*. To give each subject a score on each of these components, we standardized a subject's raw scores on each control measure and then weighted the scores by the loading of each measure on the first and second factors, respectively. These weighted loadings were summed to yield component scores for activity and anxiety for each subject.

The subjects derived from two samples as described in the *Method* section. Comparisons between these two samples indicated that there were no differences in mean performance between them on any task ($p > .05$ on each task). As noted earlier, the two samples experienced one difference in the sequencing of tasks:

The first sample experienced operant acquisition as its fourth task and fear conditioning as its last task, whereas the second sample experienced these two tasks in reverse order. The two samples produced equal mean correlations between these two tasks with the other tasks in the battery. For fear conditioning, correlations averaged −.13 and −.07 for the first and second samples, respectively. For operant acquisition, correlations with other tasks averaged −.07 and −.03, respectively. In addition, the correlations between these two tasks were .04 and .00 in the first and second samples, respectively. The two samples were combined in all analyses.

Figure 1 illustrates acquisition for the four tasks in which trial information was available. For operant acquisition, the distribution

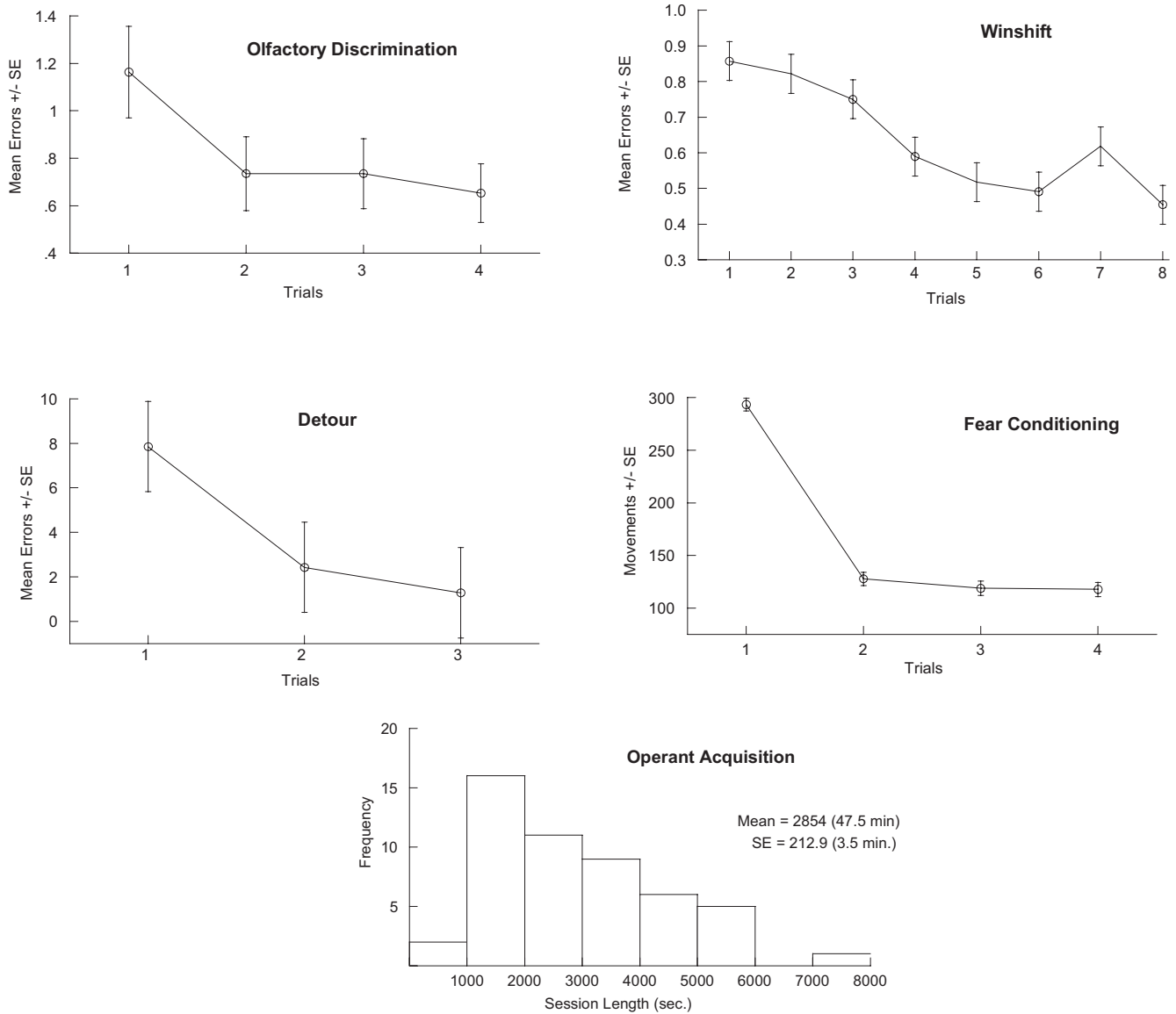*Figure 1.* Acquisition for each task in Experiment 1. For olfactory discrimination, winshift, and the detour problem, the dependent measure was mean errors per trial. For fear conditioning, movements per presentation of the conditioned stimulus was used as the dependent measure, and for operant acquisition, the time taken to reach a criterion of 25 nose-pokes constituted the dependent measure.

of the session times needed to meet the response criterion of 25 nose-pokes is presented. All but one subject reached criterion on this task within the allotted time of 90 min. Data for that one subject were not included in the figure but were included in the analyses. It can be seen that for tasks with multiple trials there were reductions in errors over the course of training: For olfactory discrimination, detour, and fear conditioning, these reductions were statistically reliable ($p < .05$; for the winshift task, $p = .08$). For operant acquisition, the distribution of session times indicated that subjects averaged 47.5 min to complete the acquisition criterion ($SE = 3.5$ min).

Table 1 gives correlations across tasks, including the two control components. The reliability of each task using Cronbach's alpha statistic is also given in parentheses. The dependent measure for the four tasks with multiple trials was the mean number of errors averaged across trials. For operant acquisition, the dependent measure was the time taken to reach the criterion of 25 nose-pokes. It can be seen in Table 1 that 5 of the 10 correlations between learning tasks were negative. The average correlation between the learning tasks was $-.03$.

Task reliabilities averaged .54. In Locurto et al.'s (2003) experiment, during which tasks were run for a number of sessions, reliabilities averaged .88 for errors. The higher reliabilities in Locurto et al.'s experiment were a function of having many more items (sessions in that study) from which to compute consistency. Cronbach's alpha is a function of the average correlation between all possible items and the total number of items in a given measure. In the present studies, items refer to trials. The difference between studies may also be partly a function of the fact that session scores, representing as they do average performance across a number of trials, tend to produce less interitem variability than do trial scores. These reliabilities may be compared with the most reliable of psychometric instruments, human cognitive batteries, that yield alpha measures of nearly .90 (Murphy & Davidshofer, 1988). Nunnally (1967) has argued that maximal consistency is found when reliabilities approach .80. The average reliabilities in this study fall below that criterion, whereas the reliabilities from Locurto et al. match or exceed that criterion. The relatively low reliabilities in the present study contributed to the relatively low average correlations observed.

In studies of individual differences in human cognition, the standard finding is that all tasks correlate positively and load positively on the first principal component. Table 2 gives two principal-components analyses without rotation, one in which the

Table 1
Correlation Matrix for Experiment 1

Task (α)	1	2	3	4	5	6	7
1. Olfactory (.70)	—						
2. Winshift (.21)	−.21†	—					
3. Detour (.39)	.27*	.08	—				
4. Fear (.85)	−.02	−.20†	−.28*	—			
5. Operant (na)	−.04	.06	.03	.00	—		
6. Activity	.18	−.03	−.28*	−.07	.03	—	
7. Anxiety	−.30*	.06	.11	.15	.04	−.30*	—

Note.   na = not applicable.
† $.05 < p < .10$.   * $p < .05$.

Table 2
Component Matrix for Experiment 1 With and Without Control Components

	Components				
	Controls included			Controls excluded	
Task	1	2	3	1	2
Olfactory	.66	.24	.54	.40	.74
Winshift	−.13	.37	−.71	.31	−.74
Detour	−.06	.84	.34	.81	.19
Fear	−.24	−.66	.37	−.71	.29
Operant	−.07	.05	−.24	.06	−.25
Activity	.68	−.33	−.30	—	—
Anxiety	−.78	−.01	.20	—	—
Eigenvalue	1.56	1.41	1.25	1.41	1.29
Percentage of variance	22	21	18	28	26

Note.   Dashes indicate not applicable.

control components were included, the other in which they were omitted. It can be seen that with control components included not all tasks loaded on the first principal component. The analysis yielded three components with eigenvalues greater than 1.0, a common standard for evaluating the significance of a component. The first component appeared to reflect the difference between the two control components rather than any of the cognitive tasks. The second component was marked principally by the differences between the detour task and fear conditioning. The third component was marked by the difference between the winshift task and the olfactory discrimination task. These components accounted, respectively, for 22%, 21%, and 18% of matrix variance. Error scores were also submitted to an oblique rotation to determine the correlation between components. No correlation exceeded .07, suggesting that these components marked essentially independent dimensions of performance.

The right-hand portion of Table 2 presents analysis with the control procedures eliminated. The first principal component in this second analysis accounted for 28% of matrix variance, and all tasks except for fear conditioning loaded positively on that component. When the fear task was eliminated in a separate analysis, the percentage of variance accounted for by the first component for the remaining four tasks rose to 33%, but the winshift task loaded negatively on this component.

Two additional analyses were conducted. In the first analysis each subject's performance on each cognitive task across trials was subjected to linear regression, and the slope for those scores was used to represent the subject's performance. This analysis used slope to reflect a subject's reduction in errors across trials irrespective of the mean number of errors made on a task, the dependent measure used in the preceding analyses. Control components were not included in this analysis. This analysis yielded the same picture of individual differences as did the analysis of error data. Three components with eigenvalues emerged that accounted for, respectively, 30%, 23%, and 21% of variance. The first component was marked principally by the difference between the fear conditioning and detour tasks.

A second alternative analysis used latencies rather than errors as the dependent measure for the olfactory, detour, and winshift tasks. The resulting principal-components analysis was similar to the analysis of errors. With control components removed, two components with eigenvalues exceeding 1.0 emerged, accounting respectively for 31% and 25% of matrix variance. The first component was marked by the winshift and detour tasks, whereas fear conditioning and operant acquisition loaded negatively on this factor.

Experiment 2

In Experiment 1, the structure of individual differences did not resemble that seen typically in studies of human cognition, wherein all correlations between tasks are positive and there is a robust first component. This conclusion was mitigated to some extent when control components were excluded in the analysis. In that case there was more positive evidence for the presence for a robust first component. To assess the generality of that conclusion, in Experiment 2 we formed a battery that included two of the tasks used in Experiment 1, the detour task and fear conditioning, along with three new tasks, a four-arm radial maze task, an olfactory foraging task, and a water escape maze task drawn from the series of tasks developed by Hebb and Williams (1946, Problem 5). Additionally, fear conditioning included two new measures, an observational measurement of freezing during conditioning and a posttest assay of lick suppression.

The changes in tasks were made for a number of reasons. The olfactory foraging task was used to replace olfactory discrimination, given that in the discrimination task there was an inflation of errors over the last two trials, suggesting perhaps a satiation effect or a loss of attention. The four-arm radial maze task introduced a land-based working memory task into the battery and replaced the operant acquisition task. The operant acquisition task appeared isolated in whatever processes it tapped in that it did not generate a high positive loading on any of the three components extracted. That same pattern of low loadings, particularly on the first component, led to the use of a Hebb–Williams maze problem in place of the winshift problem as a water escape task. The changes in fear conditioning and the inclusion of a freezing measure and a posttest lick suppression measure were undertaken to reflect measures commonly used in fear conditioning (e.g., Bouton & Bolles, 1980; Matzel et al., 2003).

Method

Subjects

Fifty-one heterogeneous stock mice obtained from the Institute for Behavior Genetics, University of Colorado, served in the experiment. They were housed and maintained as in Experiment 1. They consisted of two samples. The first sample consisted of 38 subjects; the second sample consisted of 13 subjects. As was the case in Experiment 1, these subjects were approximately 90 days old at the start of training and were within 1 week of 170 days old at the completion of training.

Apparatus, Task Descriptions, and Procedures

The battery consisted of two control procedures and five cognitive tasks. The control procedures were the open field and light–dark test used in Experiment 1.

Detour. This problem was run as in Experiment 1.

Hebb–Williams Problem 5. The original Hebb–Williams maze (1946) consisted of a series of 12 problems. Each problem partitioned an enclosed space so that a different route to the goal area was required. The start and goal areas remained the same on each problem. Our version of this task was run in a $51 \times 51 \times 16$ cm arena constructed of Plexiglas with 10 cm between adjacent alleys. Our development of a water-maze version of this task used five problems from Rabinovitch and Rosvold's (1951) standardization of Hebb–Williams Problems 1, 3, 4, 5, and 8 (see Locurto & Scanlon, 1998). We chose Problem 5, the most difficult problem from our original set of five problems, for this acquisition battery. There were four error zones defined for this problem. Errors were defined as 7.6-cm entries into an incorrect turn or into a blind portion of the maze. A Plexiglas platform 7.6 cm square and 11 cm high served as the escape platform. The platform was submerged just below the water line. Indirect lighting was provided by three 75-watt lamps (53 lux). Water was partially replaced as needed to maintain water temperature at 18 °C ± 1 °C. Following each trial, the water was whisked to remove debris and disrupt odor trails.

Practice Problem A from the Hebb–Williams series was used during the one habituation session that preceded training. Each subject received six trials of this practice problem. During habituation and training, subjects were run in subgroups of four, such that each subject received its first trial before the first subject received its second trial. On each trial, subjects were given 40 s to find the platform. If they did not find it in that time period, subjects were placed on the platform for 10 s before being removed from the arena and placed in a holding cage that was lined with dry paper. The day following training, subjects received eight trials on Problem 5.

Olfactory foraging. This task was conducted in a square arena that was originally designed for automated holeboard testing. The arena measured 26 cm on a side (TruScan, Coulbourn Instruments, Allentown, PA). The holeboard floor consisted of a square matrix of 16 holes. Each hole was 2 cm in diameter and contained a receptacle that extended approximately 1.5 cm below the floor. Nose-pokes into each hole were recorded by photobeam interruptions, and the subject's path between the holes was recorded automatically. Two days prior to habituation, subjects were weighed. Food was removed from their home cage following the 2nd day's weighing. On the 3rd day subjects were again weighed, and they were given 3 g of 20-mg food pellets in a bottle cap in their home cages to adapt them to the food reinforcer to be used during training. On the following day, they received habituation training. On the first four habituation trials, a single food pellet was placed in each of two holes chosen from the middle two columns of four holes. The holes chosen differed on each trial, so that by the end of four trials each hole had been used once. For the next four habituation trials, the same procedure was used with different pair of holes chosen from the middle two columns on each trial. In addition, on these four trials each hole was covered by a Teflon disc that had to be displaced by the animal to obtain the food pellet. Following habituation training, subjects received 3 g of food pellets in a bottle cap in their home cage.

The following day was the training session. Subjects were run through eight trials. On each trial the eight middle holes were covered with Teflon discs. Five microliters of banana extract (McCormick, Sparks, MD) were placed on top of two of the holes. For the remaining holes the Teflon discs were covered with 5 μl of tap water colored with yellow food coloring to be indistinguishable (to the eye) from the banana extract. A single food pellet was placed in the holes associated with banana extract. The dependent measure was the number of holes visited before the two food pellets were found. The holeboard and holes were cleaned with a 50% ethanol/50% water solution between trials. Subjects were given a maximum of 10 min to locate the food pellets on each trial.

Four-arm radial maze. This task was conducted in an arena that had an octagon-shaped hub and four runways, each 37 cm in length (all components from Coulbourn Instruments, Allentown, PA). The runways were constructed of clear Plexiglas sides and wire mesh floors. At one end of each runway was a liquid feeder that delivered .02 ml of skim milk. At the

other end of each runway was a guillotine door. On the day preceding feeder training, the subjects were weighed, and water was removed from their home cages. On the next day, habituation occurred. For habituation, each subject was confined in an arm that was not used during training. That arm was 11 cm in length. At the start of the session the feeder was activated, and the interior of the feeder was illuminated. Feeder entries were reinforced on a fixed-ratio one schedule. Subjects were required to make 25 nose-pokes within a 60-min period. Following this session animals received 30 min of water access in their home cages.

Testing was conducted on the next day and consisted of eight trials. Each trial began with the subject placed in the middle of the octagon hub. All four arm doors were open. A trial continued until the subject had received one reinforcer from all four arms. Following the first received reinforcement in an arm, no further reinforcements were possible in that arm. Subjects were given 30 min to complete a trial. Only one subject on one trial failed to complete the trial within this allotted time. The dependent measure was the number of arms visited in completing the trial requirement.

Fear conditioning. This task differed from the fear-conditioning task used in Experiment 1 in the dependent measures used. The protocol took 4 days to complete. On the 1st day animals were weighed, and then water was removed from their home cage. On Days 2 and 3 the subjects spent 30 min in an octagon-shaped arena (Coulbourn Instruments, Allentown, PA). One side of the octagon contained a recessed liquid drinking tube. Subjects were given 30 min each day to drink. Licks were recorded via photobeam interruptions. Following each session, subjects were given 30 min of water access in their home cages. On Day 4 fear conditioning took place in the manner described in Experiment 1: four auditory CS presentations of 30 s duration followed by a .5-s footshock at .5 mA. The ITI was 3 min. During conditioning, subjects were monitored via a video camera, and instances of freezing were recorded every 6 s during the 30 s that preceded each CS and during each CS. *Freezing* was defined as no visible movement except for breathing. A single observer scored freezing for all fear-conditioning trials. Following conditioning, subjects were returned to their home cage for 1 hr. No water access was allowed during this period. At the end of 1 hr, each subject was given the opportunity to drink in the octagon hub. Following the first 50 licks a 30-s pre-CS period was followed by a CS presentation for 30 s. Licking was recorded during the pre-CS and CS periods. Subjects ran through two trials in this manner.

Results and Discussion

As was done in Experiment 1, the six dependent measures taken from the open field and light–dark test were subjected to principal-components analyses to reduce these measures to a smaller number of components. Two components emerged with eigenvalues exceeding 1.0. These two components were similar to the two components that emerged from Experiment 1, although the pattern of loadings was slightly different. The first component, accounting for 32% of matrix variance, was marked by the difference between center crossings (loading = .86) and thigmotaxis (loading = −.81) in the open field. This component was termed *anxiety* and corresponds to the anxiety component extracted in Experiment 1. The second factor reflected the distance (loading = .73) and rearings (loading = .93) measures in the open field and accounted for 26% of variance. This measure was termed *activity*. As was the case in Experiment 1, each subject's performance on each control measure was standardized and then weighted by the component loadings to give each subject a score for anxiety and activity.

The two groups composing the overall sample did not differ on any dependent measure ($p > .05$). As a result, the combined sample was used in all analyses. Figure 2 illustrates acquisition for

the five tasks during the test session. For the four tasks in which errors were measured—olfactory foraging, detour, Hebb–Williams, and the four-arm maze—error reductions were statistically reliable across training ($p < .05$ for each task). For olfactory foraging, errors were defined as the number of holes visited in locating the two reinforcers per trial. Subjects ran through eight trials on this task. Analysis of the Trial × Trial data indicated that subjects' performance peaked at Trial 6. Performance averaged across those six trials was used in analyses. Error reductions on this measure were also reliable at $p < .05$. For fear conditioning, the changes in freezing during conditioning were used as the dependent measure to assess acquisition. There was a significant increase in freezing as a function of training ($p < .05$).

Table 3 provides the correlations between the five learning tasks and the two control components, as well as alpha coefficients. In these analyses, fear conditioning was measured as lick suppression during the test that followed fear conditioning. This lick-suppression measure did not correlate significantly with average freezing measured during fear conditioning ($r = .04$; see Bouton & Bolles, 1980, for evidence that different fear measures correlate). Substituting freezing for lick suppression in the analyses presented in Tables 3 and 4 did not alter the pattern of correlations or the loadings on the principal components. In Table 3 it can be seen that all but one of the correlations between the learning tasks were positive. The average correlation among the learning tasks was .15. The average alpha coefficient for this battery was .58.

Table 4 provides the principal-components analysis. The left-hand portion of the table provides an analysis that included control components. The right-hand portion of the table provides an analysis with the controls excluded. As was the case in Experiment 1, when controls were included three components emerged with eigenvalues greater than 1.0. In contrast to Experiment 1, in which the first component reflected the control procedures, the first component in this experiment was marked by each of the cognitive tasks except Hebb–Williams. With the control tasks included, this component accounted for 26% of matrix variance. The second component reflected the differences between the activity measure and four of the five cognitive tasks. The third component was marked principally by Hebb–Williams. These last two components accounted for 20% and 15% of variance, respectively. These data were also subjected to an oblique rotation to determine the correlation between components. The average correlation was −.03, indicating that these three components were measuring essentially independent dimensions of performance. With control components excluded, principal-components analysis produced two components with eigenvalues exceeding 1.0. The first component accounted for 34% of matrix variance, and all learning tasks loaded positively on this factor.

As was done in Experiment 1, two alternative analyses were conducted. In the first alternative analysis, the scores used in the main analysis were converted to slopes. For this analysis, fear was measured by the change in freezing over the course of the four trials of acquisition. The analysis of slopes revealed two components that each accounted for 27% of matrix variance. The first component was marked principally by the olfactory foraging and fear-conditioning tasks. In the second analysis, latencies were analyzed for each task. Inasmuch as fear conditioning did not produce a latency measure, freezing and lick suppression were used in separate analyses as the dependent measure. In both
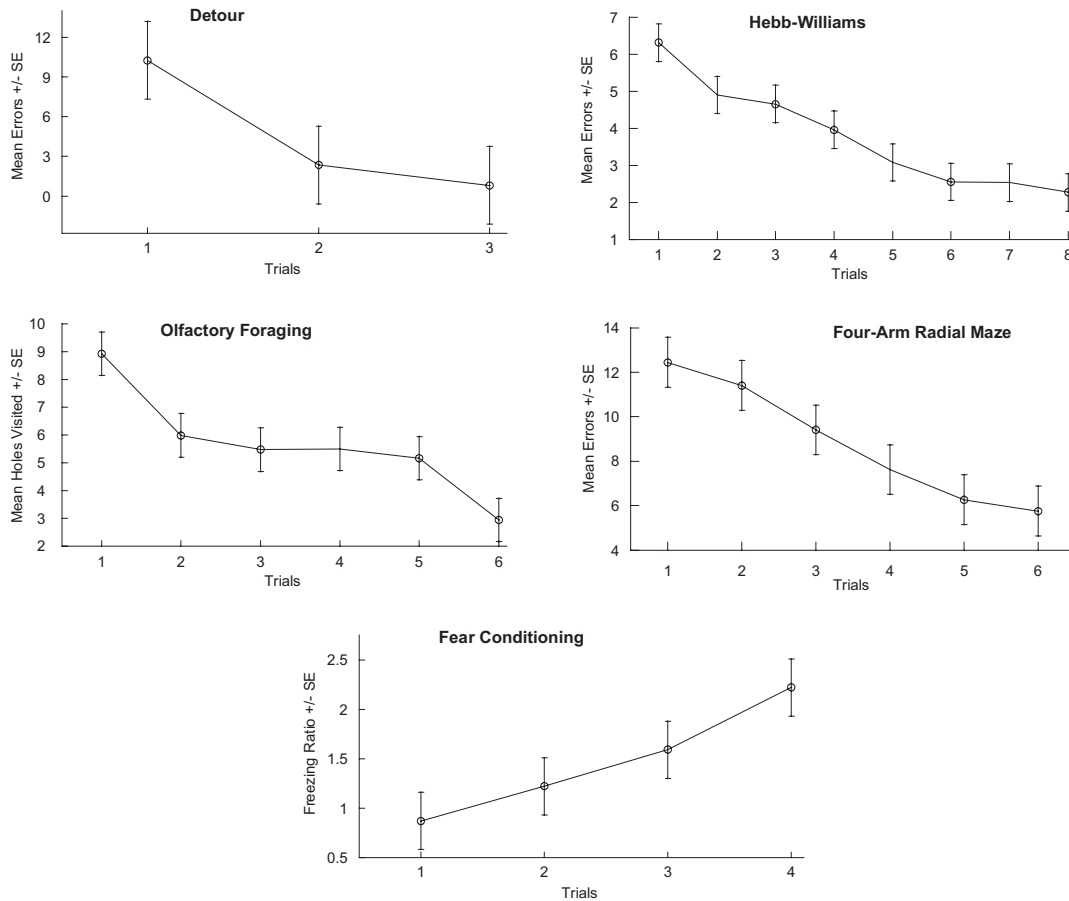
Figure 2. Acquisition for each task in Experiment 2. For the detour problem, Hebb–Williams, radial maze, and the holeboard, mean errors per trial was the dependent measure. For fear conditioning, the changes in the frequency of freezing between pre-conditioned stimulus and conditioned stimulus periods constituted the dependent measure.

analyses, all learning tasks loaded positively on the first principal component except Hebb–Williams. The first component accounted for 31% of matrix variance in each of these analyses.

General Discussion

Human cognitive batteries typically reveal the presence of a robust first principal component on which all tasks load positively.

In the present experiments, when control components were included in the analysis, the structure of individual differences did not indicate the presence of a general cognitive factor. Rather, a modular structure was suggested in which variance attached more to individual tasks than to a general component. When control procedures were eliminated, there was better evidence for the presence of a general cognitive factor, particularly in Experiment 2.

Table 3
Correlation Matrix for Experiment 2

Task (α)	1	2	3	4	5	6	7
1. Detour (.66)	—						
2. Hebb–Williams (.52)	.02	—					
3. Olfactory (.40)	.15	.02	—				
4. Radial maze (.72)	.40**	−.06	.17	—			
5. Fear conditioning (.62)	.14	.09	.30*	.25†	—		
6. Activity	−.18	−.26*	−.01	.05	.25†	—	
7. Anxiety	.17	−.06	.03	.09	−.21†	.24†	—

† $.05 < p < .10$.   * $p < .05$.   ** $p < .01$.

The finding that the structure of individual differences on cognitive tasks appears modular in the presence of control procedures is consistent with our earlier study (viz., Locurto et al., 2003) in which a battery of six tasks was continued for a number of sessions until performance approached asymptote. One difference between the two studies is that in the earlier study variance owing to activity and anxiety (termed *stress* in that earlier study) did not dominate the first component but instead appeared among the residual components. It may be that early in training differences in activity and anxiety contribute more to individual differences than they do after extended training. This idea would be consistent with Matzel et al. (2003), in which an exploration measure loaded on the first component at least as strongly as did cognitive tasks in a battery of rapid acquisition tasks.

With control components removed, the first principal component accounted for an average of 31% of the individual difference variance across experiments, and most if not all cognitive tasks loaded positively on the first component. Although this analysis may suggest the presence of a general process, it also reveals one of the problems with data reduction analyses of this sort. The outcome depends importantly on which variables (tasks) are entered into the analysis.

This fact makes it particularly difficult to compare these results with the results from human cognitive batteries with respect to the dominance of a general cognitive component. One apparent difference is that the tasks used in human cognitive batteries do not require the acquisition of new responses, as did the tasks in these batteries. It should be added, though, that in humans the processes tapped by learning tasks and tasks typically included in cognitive batteries are highly similar with respect to the presence of a general factor (Jensen, 1989). Human cognitive batteries also do not include assays of possible confounds owing to subject differences in anxiety or other noncognitive processes. Nor do they vary sensory modality and motivational state as did the present experiments.

In addition to the studies from our laboratory, there are studies from two other laboratories that have addressed the structure of individual differences in mice across a battery of cognitive tasks

Table 4

Component Matrix for Experiment 2 With and Without Control Components

| | Components | | | | |
| | Controls included | | | Controls excluded | |
Task	1	2	3	1	2
1. Detour	.57	.48	−.44	.65	−.34
2. Hebb–Williams	−.08	.54	.58	.05	.76
3. Olfactory	.50	.19	.41	.58	.32
4. Radial maze	.67	.22	−.34	.73	−.35
5. Fear conditioning	.69	−.12	.48	.64	.40
6. Activity	.28	−.81	.10	—	—
7. Anxiety	.49	−.35	−.08	—	—
Eigenvalue	1.82	1.41	1.07	1.70	1.07
Percentage of variance	26	20	15	34	21

Note. Dashes indicate not applicable.

Table 5

First Principal-Component Characteristics for Studies of Individual Differences in Mice

Study (α)	% Variance	Mean *r*
Present study, Experiment 1 (.54)		
With controls	22	
Without controls	28	−.03
Best case[a]	47	.22
Present study, Experiment 2 (.58)		
With controls	26	
Without controls	34	.15
Best case[b]	40	.19
Locurto et al. (2003) (.88)		
With controls	20	
Without controls	27	.12
Best case[c]	55	.39
Galsworthy et al. (2002)[d] (.73)		
With controls	28	
Without controls	31	.20
Galsworthy et al. (2005) (.67)[e]		
Without controls[f]	32	.12
Matzel et al. (2003)[gh] (not reported)		
With controls	25	
Without controls	38	.22

Note. Mean *r* refers to correlations among learning tasks only.
[a] Best case battery consisted of error and latency measures from the detour task, errors from the olfactory task, and latencies from the winshift task.
[b] Errors from the detour task and latencies from the olfactory, four-arm, and Hebb–Williams tasks.　[c] A four-task battery: a visual nonmatching to sample task from which errorless trials were used, and three tasks from which error measures were used, namely radial maze, Hebb–Williams, and place learning.　[d] Table 4, Galsworthy et al. (2002).　[e] Summary of two experiments.　[f] Control procedures not reported.　[g] Tables 4 and 6, Matzel et al. (2003).　[h] A study by Kolata et al. (2005) from this laboratory shows a similar matrix but was excluded from Table 5 because of a low sample size (*n* = 21).

using varied motivational and sensory conditions (viz., Galsworthy et al., 2002, 2005; Matzel et al., 2003). A summary of the primary findings from these studies, along with the results from our laboratory, is given in Table 5. For each study the percentage of variance is reported for the unrotated first principal component. The unrotated solution is preferred by those who wish to maximize the percent variance accounted for by the first component, because that component represents the best fit to the matrix of correlations. The average correlation between the tasks for each study is also given, as well as reliability estimates when available.

Table 5 also provides three different data-reduction outcomes for the three experiments from our laboratory, one with control components included, the second without the control components, and a third analysis that is termed *best case*. This last analysis was undertaken because one of the differences between our work and studies from other laboratories is in the choice of reported measures. In Locurto et al.'s (2003) study, we used an aggregated measure that represented errors, latencies, and errorless trials in a single measure. In the present set of experiments, we used errors as our fundamental measure whenever possible. In Galsworthy et al.'s (2002) study and Matzel et al.'s (2003) study, errors were used for some tasks and latencies for others. Furthermore, in Galsworthy et al.'s study, eight measures were reported from five tasks. For three of the arenas multiple measures were reported, for

example, latencies and errors for Hebb–Williams were both used in the principal-components analysis.

The use of multiple measures from the same task as well as the use of different dependent measures across tasks may represent a best case picture of the strength of the first component. This practice is similar in some respects to the method commonly followed in the construction of human cognitive batteries in which the presence of a robust general factor is assumed. Tasks are chosen that load on that factor, whereas other tasks are eliminated because they do not do so (Anastasi, 1989; Mackintosh, 1998). The best case measures reported in Table 5 for our studies stem from that type of analysis, one in which the set of measures that produced the strongest first component is reported.

The data in Table 5 illustrate commonalities across the studies in the three laboratories. With control procedures included, the average variance accounted for by the first principal component in experiments from our laboratory was 23%. For the studies by Galsworthy et al. (2002) and Matzel et al. (2003), that average was 26.5%. Excluding control procedures raised the variance accounted for in our studies to 30%. In the other studies that average rose to 34%. Our average correlation across learning/cognitive tasks was .10, whereas the average correlation for the other studies was .18. But when we formed best case batteries using either latencies or errors to produce the strongest first component, our average correlation rose to .30, and the variance accounted for rose to 47%.

The commonalities between these studies suggest that the magnitude of any general factor is similar across these studies. Comparisons between studies with respect to reliabilities are more difficult to judge. Reliabilities were not reported by Matzel et al. (2003), and Galsworthy et al. (2002) reported reliabilities only for tasks that were run for multiple sessions. The average correlation reported in Table 5 for that study would likely be lower if reliabilities for single session tasks had been included.

Galsworthy et al. (2002, 2005) and Matzel et al. (2003) drew the conclusion that there was evidence for a general learning or cognitive ability. In doing so they emphasized the matrix of positive correlations and the fact that all tasks loaded positively on their first component. We noted that same pattern in Experiment 2 when control components were eliminated, both for errors and for latencies. Experiment 1 offered a similar picture when control components were eliminated. In that experiment only the winshift task loaded negatively on the first component, but the variance accounted for by the first component was 28%, quite close to the accounted-for variance average in the other studies.

Whichever features of these studies are emphasized, it is clear that the characteristics of the first component are more modest in these studies using mice than is typically observed in standardized human cognitive batteries. It is often the case that the first component in human cognitive batteries accounts for more variance than the sum of all other components combined (Jensen, 1998). As an example, the correlation matrix for the Wechsler Adult Intelligence Scale—Revised consists of 11 tasks (Mackintosh, 1998, Table 2.2). These tasks produced a matrix of 55 correlations, all of which were positive. The average correlation was .50 (range = .33–.81). As is true for other versions of this test, the first factor accounted for more than 50% of matrix variance (Wechsler, 1981). The only data point from mice that approaches the magnitude of these data comes from the best case battery from Locurto et al.

(2003) in which the first component accounted for 55% of matrix variance, and the average correlation was .39.

It should be pointed out that the Wechsler Adult Intelligence Scale—Revised was constructed in the manner of a best case scenario, in that tasks were included in the final battery only if they correlated positively with other tasks and loaded positively on the first component. That is, the presence of general cognitive factor (g) was assumed, and tasks were chosen that verified its presence. An analogous best case procedure may well be justified in animal studies as a way of approaching the kind of test-construction process that attends human cognitive batteries. As noted earlier, studies of individual differences in mice have also used a number of control procedures to assess the noncognitive components of performance. Nothing similar is typically done in human cognition. In addition, for the studies cited in Table 5, care was taken to ensure that tasks tapped a variety of behavioral systems, meaning the sensory domain, response, motivation, and reinforcer used for each task. These variations were instituted so as to reduce the possibility of artifacts producing the observed correlations. In human cognitive batteries, tasks are for the most part administered in the same manner, with verbal instructions attendant to each task. Some batteries are typically administered as paper-and-pencil tasks, particularly those that are designed for group administration. (See Anastasi, 1989, for a description of these different batteries.) There is little if any effort to vary the test-taking situation, motivation, or sensory domains tapped by the battery.

In this sense a typical human cognitive battery may be more like a series of tasks administered to mice in which the motivational system and other aspects of the procedures remain similar across tasks. Our laboratory's first work in the area of individual differences in mice used that type of battery (viz., Locurto & Scanlon, 1998). This study was not included in Table 5 because it did not vary motivation across tasks, as did the other studies cited in the table. All tasks were run under water escape motivation, and although the arenas differed between tasks, all tasks required either spatial or visual cues for solution. A reanalysis of Locurto and Scanlon's (1998) results using the best case scenario approach yielded a battery of seven measures taken from five tasks (Morris maze transfer task, Hebb–Williams, place learning, spatial discrimination, and visual discrimination). The 21 correlations were all positive, with an average correlation of .31. The first component accounted for 43% of matrix variance. These characteristics approach the findings typically reported for human cognitive batteries.

It should be recognized that there is a problem attendant to the use of best case procedures. If analysis begins with the assumption of a robust general component, any task or measure is justified on the basis of its loading on the first component. In the process of constructing the battery in this manner, one is then also constructing g. In doing so there may be no fundamentally sound a priori rationale for task or measure selection. Why should some tasks be included and others not included? Why should latencies from one task and errors from another be chosen? If there is no rationale other than the predetermined outcome, one may be ignoring important exceptions to the idea that most of the variance is captured by a general component. In our work we have reported all tasks that were used in a particular battery irrespective of how they behaved relative to other tasks or of how they loaded on the first component. We can, by a best case procedure, construct a battery

that evidences a strong first principal component. If that result were all that was reported, we would be ignoring palpable evidence that many tasks correlate only modestly and that a full rendering of individual difference structure via principal components often reveals a picture with numerous small but significant components rather than one robust general component.

This same oversight may be true for human cognitive batteries. How many tasks are discarded to support a robust g? Do those eliminated tasks tap processes that might reasonably be regarded as important components of human cognition, irrespective of their loadings on the first component? These questions are not answerable given the kinds of task- and battery-construction procedures that commonly attend the finding of a strong general process.

References

Anastasi, A. (1989). *Psychological testing* (7th ed.). New York: Macmillan.

Anderson, B. (1993). Evidence from the rat for a general factor that underlies cognitive performance and that relates to brain size: Intelligence? *Neuroscience Letters, 153,* 98–112.

Barlow, J., Cosmides, I., & Tooby, J. (Eds.). (1992). *The adapted mind: Evolutionary psychology and the generation of culture.* New York: Oxford University Press.

Bouton, M. E., & Bolles, R. C. (1980). Conditioned fear assessed by freezing and by the suppression of three different baselines. *Animal Learning & Behavior, 8,* 439–454.

Collias, N. E., & Collias, E. C. (2004). Comparison of vocal signals of three species of African finches. *Behavior, 141,* 1151–1171.

Gallistel, C. R. (2003). The principle of adaptive specialization as it applies to learning and memory. In R. H. Kluwe & G. Kyer (Eds.), *Principles of learning and memory* (pp. 259–280). Cambridge, MA: Birkhauser.

Galsworthy, M. J., Paya-Cano, J. L., Monleon, S., Gregoryan, G., Fernandes, C., Schalkwyk, L. C., & Plomin, R. (2005). Assessing reliability, heritability and general cognitive ability in a battery of cognitive tasks for laboratory mice. *Behavior Genetics, 35,* 675–692.

Galsworthy, M. J., Paya-Cano, J. L., Monleon, S., & Plomin, R. (2002). Evidence for general cognitive ability (g) in heterogeneous stock mice and an analysis of potential confounds. *Genes, Brain and Behavior, 1,* 88–95.

Hebb, D. O., & Williams, K. (1946). A method of rating animal intelligence. *Journal of Genetic Psychology, 34,* 59–65.

Jakubowska, D. E., Gumusbas, U., & Kara, F. (2003). Individual variation in the spatial reference and working memory assessed under allothetic and idiothetic orientation cues in rat. *Acta Neurobiologiae Experimentalis, 63,* 17–23.

Jensen, A. (1989). Learning and intelligence. *Learning and Individual Differences, 1,* 37–62.

Jensen, A. (1998). *The g factor: The science of mental ability.* New York: Praeger.

Kolata, S., Light, K., Townsend, D. A., Hale, G., Grossman, H. C., & Matzel, L. D. (2005). Variations in working memory capacity predict individual differences in general learning abilities among genetically diverse mice. *Neurobiology of Learning and Memory, 84,* 241–246.

Landeira-Fernandez, J., Fanselow, M. S., DeCola, J. P., & Kim, J. J. (1993). Naltrexone does not disrupt acquisition and performance of inhibitory conditioning. *Bulletin of the Psychonomic Society, 31,* 591–594.

Locurto, C. (1997). On the comparative generality of g. In W. Tomic & J. Kigma (Eds.), *Advances in cognition and education: Vol. 4. Reflections on the concept of intelligence* (pp. 79–100). Greenwich, CT: JAI Press.

Locurto, C., Fortin, E., & Sullivan, R. (2003). The structure of individual differences in heterogeneous strain mice across problem types and motivational conditions. *Genes, Brain, and Behavior, 2,* 1–16.

Locurto, C., & Scanlon, C. (1998). Individual differences and a spatial learning factor in two strains of mice. *Journal of Comparative Psychology, 112,* 344–352.

Mackintosh, N. J. (1998). *IQ and human intelligence.* Oxford, England: Oxford University Press.

Macphail, E. M. (1987). The comparative psychology of intelligence. *Behavioral and Brain Sciences, 10,* 645–695.

Matzel, L. D., Han, Y. R., Grossman, H., Karnik, M. S., Patel, D., Scott, N., et al. (2003). Individual differences in the expression of a "general" learning ability in mice. *Journal of Neuroscience, 23,* 6423–6433.

McClearn, G. E., Wilson, J. R., & Meredith, W. (1970). The use of isogenic and heterogenic mouse stocks in behavioral research. In G. Lindzey & D. Thiessen (Eds.), *Contributions to behavior genetic analysis: The mouse as a prototype* (pp. 3–22). New York: Appleton-Century-Crofts.

Mihalick, S. M., Langlois, J. C., Krienke, J. D., & Dube, W. V. (2000). An olfactory discrimination procedure for mice. *Journal of the Experimental Analysis of Behavior, 73,* 305–318.

Miller, L. T., & Vernon, P. A. (1992). The general factor in short-term memory, intelligence, and reaction time. *Intelligence, 16,* 5–30.

Murphy, K. R., & Davidshofer, C. O. (1988). *Psychological testing: Principles and applications.* Englewood Cliffs, NJ: Prentice Hall.

Nikkhah, G., Rosenthal, C., Hedrich, H. J., & Samii, M. (1998). Differences in acquisition and full performance in skilled forelimb use as measured by the "staircase test" in five rat strains. *Behavioural Brain Research, 92,* 85–95.

Nunnally, J. (1967). *Psychometric theory.* New York: McGraw-Hill.

Rabinovitch, M. S., & Rosvold, H. E. (1951). A closed-field intelligence test for rats. *Canadian Journal of Psychology, 5,* 122–128.

Seligman, M. E. P., & Hager, J. L. (Eds.). (1972). *Biological boundaries of learning.* New York: Appleton-Century-Crofts.

Tooby, J., & Cosmides, I. (2000). Toward mapping the evolved functional organization of mind and brain. In M. S. Gazzaniga (Ed.), *The new cognitive neurosciences.* Cambridge, MA: MIT Press.

Verrico, C. D., Jentsch, D. J., Roth, R. H., & Taylor, J. R. (2004). Repeated, intermittent Delta–sup-9-tetrahydrocannabinol administration to rats impairs acquisition and performance of a test of visuospatial divided attention. *Neuropsychopharmacology, 29,* 522–529.

Wechsler, D. (1981). *The WAIS–R manual: Wechsler Adult Intelligence Scale—Revised.* New York: Harcourt Brace Jovanovich.