Behavioral Neuroscience
2015, Vol. 129, No. 4, 533–539

0735-7044/15/$12.00 http://dx.doi.org/10.1037/bne0000078

BEHAVIORAL NEUROSCIENCE AT 30

Using the Spatial Learning Index to Evaluate Performance on the Water Maze

Inês Tomás Pereira and Rebecca D. Burwell
Brown University

The Morris water maze was developed in 1981 and quickly became the standard task for assessing spatial memory and spatial navigation. Twenty years ago, Gallagher, Burwell, and Burchinal (1993) reported new variables and measures, including a spatial learning index, that greatly enhanced the utility of the Morris water maze for assessing subtle differences in performance on the task. The learning index provided a single number that could be used to elucidate neurobiological measures of hippocampal dysfunction, for example, correlation of learning performance with a biomarker of aging. In this review, as part of the commemoration of the 30th anniversary of *Behavioral Neuroscience,* we describe how the spatial learning index has contributed to the field of learning and memory, how it has advanced our understanding of normal and pathological cognitive aging, and how it has contributed to translation of findings into other species. Finally, we provide instruction into how the learning index can be extended to other tasks and data sets.

Keywords: proximity, aging, learning, behavior, rat

The Morris water maze task was described more than three decades ago (Morris, 1981; Morris, Garrud, Rawlins, & O'Keefe, 1982) and revolutionized the way neuroscientists study navigation and hippocampal-dependent memory. The article under review (Gallagher, Burwell, & Burchinal, 1993) introduced a new measure, proximity to platform location, and a method for condensing a complex behavioral task into a single index of performance. This spatial learning index greatly facilitates within-group comparisons and correlations with neurobiological markers or other behavioral measures. Additionally, the 1993 study demonstrated that, similar to humans, the Long-Evans rat model displays wide variability in age-related cognitive decline. The proximity measure proved to be sensitive to subtle changes in behavior among aged individuals, clearly revealing the increased variability in aged rats, compared to young controls, on performance in the water maze. For many studies, this increased variability in a single measure was used to divide aged rats into two subgroups: one with performance comparable to young animals (and substantial sparing of cognitive function) and the other characterized by cognitive decline. The comparison between these groups of animals has contributed to identifying mechanisms that contribute to cognitive decline and those that contribute to the maintenance of function in normal aging. The development of the spatial learning index was, therefore, crucial for the advancement of the study of normal cognitive aging and continues to yield insights about the mechanisms that contribute to the maintenance of cognitive function in older age.

Originally, the Morris water maze was developed to take advantage of the innate spatial navigation and swimming abilities of rats. The task is conducted in a large tank filled with opaque water that has an escape platform concealed under water. In the spatial reference memory version, animals are placed in the tank in a different starting point for each trial and have to use the spatial layout of the room to navigate to the platform that is kept in a constant location. Typically, normal young animals learn to swim to the platform more quickly over time and reach behavioral asymptote over several trials. These learning trials, sometimes called training trials, can be spaced out over days or condensed into one day. Additionally, probe trials may be used at different time points during learning to assess the pattern of spatial memory acquisition. For these probe trials, the platform is made temporarily unavailable, and the search pattern of the animals is analyzed with respect to the position of the platform in previous learning trials. In recent years, taking advantage of developments in virtual machine software, some studies have adapted this task for use in human studies (Daugherty et al., 2014; Goodrich-Hunsaker, Liv-

Editor's Note. John F. Disterhoft served as action editor for this article.—RDB

Inês Tomás Pereira and Rebecca D. Burwell, Department of Cognitive, Linguistic and Psychological Sciences, Brown University.

We thank Peter Rapp and Michela Gallagher for comments on this article. We also acknowledge the support of the NSF (IOB-1146334), DARPA (N66001-10-C-2010), the Brown Institute for Brain Sciences, and the Brown University Center for Vision Research.

Correspondence concerning this article should be addressed to Rebecca D. Burwell, Department of Cognitive, Linguistic and Psychological Sciences, Brown University, 190 Thayer Street, Box 1821, Providence, RI 02912. E-mail: rebecca_burwell@brown.edu

ingstone, Skelton, & Hopkins, 2010; Sandstrom, Kaufman, & Huettel, 1998), greatly increasing the translation potential of the findings in rodents.

Measures for Evaluating Water Maze Performance

One of the first measures used to quantify behavior in the Morris water maze task was the latency to reach the hidden platform. Although informative, this measure can vary for reasons unrelated to the learning ability of the animal. For example, different swimming speeds can lead to different latencies, even if the path taken is the same. To control for this possibility, the path length, a measure of the total distance navigated to reach the goal, was used along with latency. This measure, however, is also limited because some navigation strategies can lead to shorter path lengths despite lack of knowledge of the platform location. For example, a rat can learn that the platform is a certain distance from the wall of the maze, swim a circular path around the edge (thygmotaxis), and find the platform relatively efficiently, even if the animal has no representation of its location (see Figure 1 in Gallagher et al., 1993).

Other frequently used measures are based on probe trials. For these trials, the submerged platform is not present. Probe trials provide insight into the search strategy used by the animals to locate the missing platform and can be used to determine the nature of deficits. One drawback is that probe trials can lead to extinction learning or learned helplessness. To prevent extinction learning, some investigators have shortened the duration of the probe trial. Others have utilized a platform that can be lowered for most of the probe trial and then raised to its original location at the end, reinforcing earlier learning trials (e.g., Baxter, Bucci, Gorman, Wiley, & Gallagher, 1995). The use of an adjustable-height platform also minimizes the behavior shown by some rats of searching away from the location of the platform after repeatedly not finding it in the target location. The probe trial measures include the relative time spent in the target quadrant (or zone around the platform) and number of platform location crossings. These measures are more representative of a targeted location search, but they suffer from lack of reliability. For example, two rats might spend the same time in the target quadrant, but one may search more closely to the actual location of the platform. Likewise, a rat may search in a tight circle very close to the platform location while failing to cross the location. Additionally, for these measures, there is insufficient within-group variability to provide the parametric space necessary for reliable correlation analysis with other behavioral or neurobiological markers.

To better characterize the search strategy the animal employs in learning trials, Gallagher et al. (1993) introduced the *proximity* measure, a measure that relies on the ongoing assessment of the distance between the platform location and the animal. In the original article, proximity was calculated 10 times per second, and a mean was taken every second. The dotted lines in Figure 1 represent the 1-s mean proximity between the rat and the target platform location. From this measure, two additional variables can be derived. In learning or training trials, a *cumulative search error* is calculated by summing the 1-s proximity means for each second the rat is searching for the platform. Figure 1A illustrates how cumulative search error is computed from proximity in example trials early (left) and late (right) in learning. In these examples, the path length and latency are both greater for the early trial. Cumulative search error, the accumulation of 1-s mean proximities represented by the dotted lines, provides a more sensitive measure of spatial memory, yielding 2,388 cm for the early trial and 220 cm for the late trial. The distance between the entry point into the maze and the location of the platform is sometimes subtracted from the cumulative search error to normalize over different start locations.

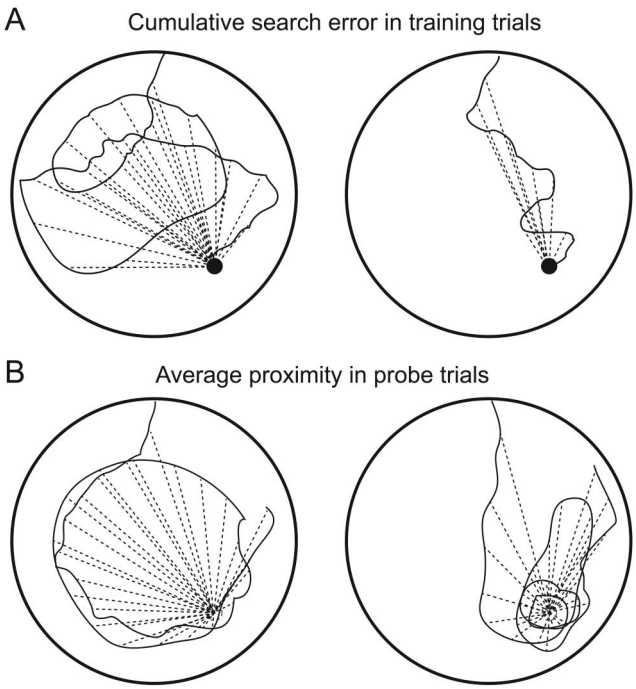

Figure 1. Examples of training trials and probe trials used for calculating measures based on proximity to the target platform location. Solid lines represent the path to the platform location. Dotted lines represent 1-s averages of the proximity of the rat to the platform location. (A) On the left, path length was 745 cm and latency was 27 s. Thus, proximity averaged for each second was summed to a cumulative search error of 2,388 cm. On the right, path length was 220 cm and latency was 9 s. Thus, proximity averaged for each of 9 s was summed to a cumulative search error of 605 cm. Cumulative search error is sensitive to both latency and path accuracy. (B) These two examples of 30-s probe trials illustrate the issues with the path length and platform crossing measures. The path shown on the right clearly exhibits a better knowledge of platform location, yet the path lengths are similar. The path length on the left is 733 cm and crosses the platform once. The path on the right is 736 cm and does not cross the platform location. Average proximity (the mean of 30 one-second averages) does distinguish performance; average proximity is 72.6 cm for the example on the left and 38.6 cm for the example on the right.

The proximity measure is even more powerful in probe trials, in which the platform is removed, especially compared to the time in quadrant and platform crossings measures (Maei, Zaslavsky, Teixeira, & Frankland, 2009). Figure 1B shows two examples of 30-s probe trials in which swimming speed and path length were comparable. In the trial on the left, the rat has adopted a strategy of swimming at a relatively fixed distance from the wall. The subject spent 10.6 s in the target quadrant and crossed the platform location one time. In the trial on the right, the rat is making a

targeted search, spending 18.9 s in the target quadrant but not actually crossing the platform location. *Average proximity* is much more sensitive for distinguishing the performances on these two example probe trials than time in quadrant or platform location crossings. In Figure 1B, again, each dotted line represents a 1-s mean proximity to the goal for each second of the 30-s probe trial. When these 30 numbers are averaged to obtain the mean proximity, the left trial has an average proximity of 72.6 cm, whereas the trial on the right is 38.6 cm. The proximity measure can also be blocked during probe trials to assess whether subjects alter their search strategy as the probe trial progresses. Such an analysis could be used to determine the optimal duration for probe trials. Thus, the sensitivity of these measures allows for better quantification of even small differences in behavior.

One interesting study performed a meta-analysis comparing the different measures used to quantify performance on the water maze (Maei et al., 2009). The data were all collected from the same laboratory, using the same water maze protocol, but the experimental groups differed, including studies of pharmacological agents, genetic models, and anatomical lesions. This data set included 1,600 individual probe trials, allowing the authors to perform Monte Carlo simulation experiments testing several conditions such as different sample sizes, different effect sizes, and the use of parametric or nonparametric statistical analyses. The water maze performance measures analyzed included platform crossings, time in target quadrant, time in target zone, and proximity. For all conditions tested, this study showed that the proximity measure is the most sensitive for detecting group differences. Importantly, this result means that the use of proximity measures can reduce the numbers of animals needed in experiments. Furthermore, in behavioral characterization facilities specializing in the characterization of mouse models, the standardized use of the proximity measure may increase throughput.

Indexing Spatial Memory in the Morris Water Maze

The proximity measures, in general, allowed for better quantification of the navigation ability of animals in the Morris water maze task. Relating performance on this task with other behaviors or with brain biomarkers was still difficult to do because the learning rate was described by repeated measures, and the asymptotic performance was often insensitive to individual or group differences. In the original article, Gallagher et al. (1993) developed the *learning index,* an additional measure that can be used for the purpose of relating spatial learning ability with other behavioral or neurobiological measures. Rats were given three trials per day. Every other day, the third trial was a probe trial in which the platform was either removed or lowered to the bottom of the pool for the initial 30 s. To prevent extinction learning, later studies adopted a submergible platform that could be raised at the location at the end of the probe trial, reinforcing earlier learning trials (Baxter et al., 1995). By the fourth probe trial, the control group had reached asymptotic performance. The average proximities for the four probe trials were combined into the learning index. Importantly, each probe trial was weighted differently. The weightings were designed to enhance the contribution of early learning, so that better performance on earlier probe trials is reflected in a better learning index. Therefore, differences in scores on early probe trials have a greater impact on the learning index than differences in the later probe trials, when it is more likely that all animals have learned the task. Generally, the procedure for calculating the weighting factors is based on performance of a cohort of control animals or on archival data. A detailed description of the method used to calculate the learning index for the water maze and other tasks is given in the section "Using a Learning Index for Other Learning Tasks."

Use of the New Measures in the Field

Since the publication of the target article, the measures developed by Gallagher and colleagues have been used in many studies to characterize water maze performance of both rats and mice. In particular, in mutant and transgenic mouse models, cumulative search error and average proximity are now frequently used to assess hippocampal-dependent memory, similar to how long-term potentiation (LTP) and long-term depression (LTD) are used to measure synaptic plasticity. For example, studies using the proximity measure established important roles for calcineurin (Malleret et al., 2001), EphB2 (Grunwald et al., 2001), CREB (Pittenger et al., 2002), CamKIIalpha (Miller et al., 2002), and NF1 and Ras (Costa et al., 2002) in spatial memory. The proximity measure has also been used in water maze procedures in which only one probe trial is used. These studies have, for example, characterized changes in synaptic plasticity and place field properties in aged animals (Barnes, Suster, Shen, & McNaughton, 1997; Rosenzweig, Rao, McNaughton, & Barnes, 1997). We propose that the proximity measure and the cumulative search error should become standard for quantification of behavior in the water maze task. This may have been difficult or cumbersome in the past due to the necessity of using methods that allow for the precise measuring of the distance between the animal and the platform several times per second. Automated video tracking software, however, has become more reliable and now often includes automated calculation of the proximity measure.

The learning index has also been used to assess hippocampal integrity, particularly in studies in which impairment is more subtle, for example, studies of cognitive aging. In an early study, the learning index was used to show that hippocampal neuron number does not predict age-related spatial memory impairment (Rapp & Gallagher, 1996). The index was also important in identifying features that positively contribute to the maintenance of cognitive abilities with aging, as is the case with the levels of hippocampal synaptophysin (Smith, Adams, Gallagher, Morrison, & Rapp, 2000). In addition, different types of hippocampal LTP and LTD are differently correlated with learning index in young and aged rat cohorts (Lee, Min, Gallagher, & Kirkwood, 2005; Yang et al., 2013). The distribution of NMDA receptor subunits is also altered in the hippocampus and prefrontal cortex of mice in a way that relates to spatial memory performance as assessed by the learning index (Magnusson, Scruggs, Zhao, & Hammersmark, 2007). The changes in glutamate receptor composition in aged animals may contribute to the physiological changes demonstrated by altered place cell properties (Barnes et al., 1997) that correlate with the learning index (Wilson et al., 2003). Additionally, subregional hippocampal decreases in the number of somatostatin- and GAD-67–expressing interneurons are related to the impairment described by the learning index (Spiegel, Koh, Vogt, Rapp, & Gallagher, 2013). Finally, changes in epigenetic mechanisms are proposed to contribute to age-related memory impairments in mice (Peleg et al., 2010), but when related to a more precise cognitive measure such as the learning index in rats, the results suggest a more

nuanced contribution of epigenetic factors to age-related cognitive impairments (Castellano et al., 2012; Tomás Pereira et al., 2013).

The learning index can also be used to relate age-related changes in performance on the water maze task to performance in other behavioral assessments, whether hippocampal dependent or not. Predictably, tasks that depend on hippocampal integrity, such as recognition memory tasks, show impairments that are correlated with spatial learning and the learning index (Robitsek, Fortin, Koh, Gallagher, & Eichenbaum, 2008). This study further demonstrates that the significant correlation is restricted to the hippocampal-based recollection process but not the familiarity component of recognition memory. Additionally, the learning index has been useful in comparisons between hippocampal impairment and declines in other cognitive domains. Although performance in an odor discrimination task is correlated with water maze performance (LaSarge et al., 2007), executive functions dependent on prefrontal cortical function, such as reversal and set-shifting, seem to show deficits that are unrelated to the learning index (Beas, Setlow, & Bizon, 2013; Schoenbaum, Nugent, Saddoris, & Gallagher, 2002). Likewise, age-related changes in reaction time are not related to spatial learning in the water maze assessed by the learning index (Burwell & Gallagher, 1993). In summary, the learning index has been used in many studies characterizing the variability of the aging process and the mechanisms that may underlie successful cognitive aging, thus suggesting avenues of intervention that may reverse those deficits.

Most studies employing the learning index to characterize performance on the water maze task (only a few of which have been reviewed here) have focused on the study of aging populations. The learning index can, however, be used to compare discrete changes in behavior in other settings. For example, the learning index was used to demonstrate that rats with selective cholinergic basal forebrain lesions show similar spatial learning and navigation properties to their control counterparts (Baxter et al., 1995). Another study used the learning index to relate water maze performance of animals with perirhinal, postrhinal, and entorhinal cortex lesions to other behavioral dimensions, such as contextual fear discrimination and passive avoidance (Burwell, Saddoris, Bucci, & Wiig, 2004).

As detailed above, the learning index has proven to be a robust and useful measure in many studies. Nevertheless, it is a reductionist measure in the sense that it reduces complex behavior into

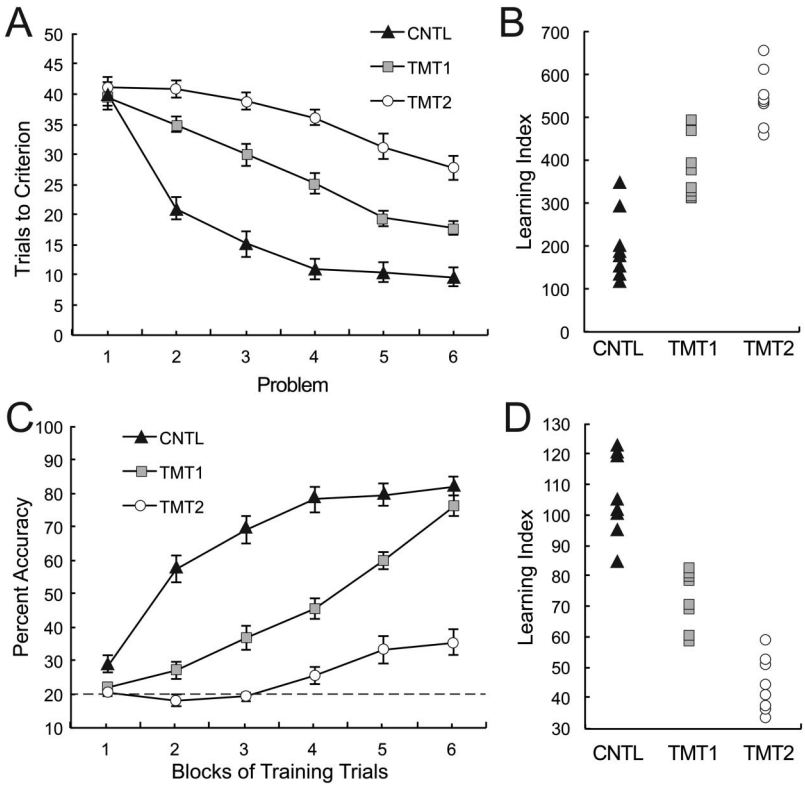

Figure 2. Test data sets illustrating how a learning index can be derived for any repeated measure that changes with learning. (A) A data set of trials to criterion (TTC) on six discrimination problems was generated for a control group (CNTL, $n = 8$) and two treatment groups (TMT1 and TMT2, $n = 8$). In this test data set, all groups show decreased TTC with each problem. TTC decreases more rapidly and is lower on the last problem for the control group. (B) Learning indices were calculated for each subject using the multipliers shown in Table 1. Indices are lower for better learners. (C) A data set of percent accuracy (chance is 20%) on a learning task was generated for six blocks of four training trials, also for three groups ($n = 8$ per group). (D) Learning indices for the subjects in C. Because the primary variable (accuracy) increases with learning, the indices are higher for better learners.

a single number. For some studies, this may not be desirable, and analysis of repeated measures of differences in behavior may be more appropriate. Another issue is that the proximity measure is not always included in automated software packages, and investigators may need to calculate the measure from x, y coordinates of the moving subject and the stationary target location. Finally, the use of proximity to calculate a learning index may not be straightforward. In the following section, we clarify the construction and use of learning indices.

Using a Learning Index for Other Learning Tasks

The ability to describe performance on a task with a single number is useful in behavioral neuroscience for reasons already described. Most learning variables, however, are collected over multiple time points. Mean performance or asymptotic performance can be used, but learning curves sometimes show that subjects often differ only in initial performance and rate of acquisition, such that mean or asymptotic performance obscures group differences. A learning index that incorporates rate of acquisition can be derived empirically from any repeated variable. In the case of Gallagher et al. (1993), the weights used in the learning index were derived from normative data collected over a large number of young rats. Many laboratories that consistently use a paradigm for multiple studies will have such normative control data. When normative data are not available, however, it is also possible to use the mean data from the control group in the study itself. Examples shown in Figure 2 are test data generated for two model experiments to illustrate how an index can be derived in a single study.

In the first example (Figure 2A), the test data are trials to criterion (TTC) for six discrimination problems. Like probe trial average proximity in the Morris water maze, TTC decreases with each problem. Such decreases in TTC might reflect the acquisition of a learning set that can be used for learning each discrimination more quickly. In this test data set, there are three groups of subjects: the group with the best acquisition is the control group, and the other two groups might be two different treatment groups. The computation of the learning index was empirically derived from the control group, and the mean control data were used to derive a set of weights, or multipliers, that favor rapid acquisition of the learning set. The multiplier for each problem was the quotient of the mean TTC in the control group for the first problem and the mean TTC for that problem, yielding a set of six multipliers shown in Table 1. TTC of the control group on the first problem is used as the numerator because this number reflects performance prior to formation of a learning set. The learning index is calculated as follows:

$$\text{Learning Index} = \sum_{i=1}^{n} (\text{Mi} \times \text{Si}),$$

where M is the multiplier, S is the score, and n is the number of scores. Because there was no difference in performance among the groups on the first problem, and presumably no learning set had been formed, TTC on the first problem does not contribute to the learning index. Here, we substituted the first multiplier with zero. Figure 2B shows the indices for subjects in each group. The distribution of individual values is still reflective of overall group differences, but in addition, each value represents, in a single number, the learning dynamics over the course of several days of training for an individual animal. For the purposes of relating learning ability to other relevant parameters (behavioral or neurobiological), the learning index provides major advantages, including better characterization of within-group differences.

The second example (Figure 2C) is based on a variable that increases with learning—in this case, percent accuracy on six blocks of four trials, in a task in which chance performance is 20%. Again, there are three groups of subjects—a control group and two treatment groups—and the computation of the learning index was empirically derived from the control group. In this case, because the number in the numerator should reflect a time point prior to any learning, accuracy for the control group on the first trial (instead of first block) was used. The multiplier for each block was the quotient of the mean accuracy of that block and the mean control accuracy on the first trial, yielding a set of six multipliers shown in Table 2. Contrary to the example shown in Figure 2A and 2B, there was learning in the first

Table 1
Multipliers for Learning Indices in Figure 2A and 2B

Problem	TTC	Multiplier formula	Multiplier value
1	40.0	M1 = C1/C1	1.0[a]
2	21.1	M2 = C1/C2	1.9
3	15.1	M3 = C1/C3	2.6
4	11.0	M4 = C1/C4	3.6
5	10.4	M5 = C1/C5	3.8
6	9.6	M6 = C1/C6	4.2

Note. Numbers 1–6 represent the number of C or M. The multiplier is the mean TTC for the first problem for the control group (in which there is presumably no learning) divided by the mean TTC for that problem. The learning index is the sum of the products of TTC and the appropriate multipliers for all problems. These multipliers are used for the control group and the treatment groups. TTC = trials to criterion; C = mean TTC; M = multiplier.
[a] The first score can be dropped because there were no group differences at this point. This can be accomplished by substituting zero for the multiplier.

Table 2
Multipliers for Learning Indices in Figure 2C and 2D

Block	Accuracy	Multiplier formula	Multiplier value
1	28.8	M1 = T1/B1	0.62
2	58.0	M2 = T1/B2	0.31
3	69.8	M3 = T1/B3	0.25
4	78.8	M4 = T1/B4	0.23
5	79.9	M5 = T1/B5	0.22
6	81.5	M6 = T1/B6	0.21

Note. The numbers 1–6 represent the number of B, M, or T. For accuracy or any variable that is averaged across a block of trials or sessions, the numerator for the multiplier is the first trial (or first session), because presumably little learning has occurred. The denominator is the average accuracy for that block. Again, the learning index is the sum of the products of the trials to criterion and the appropriate multipliers for all problems, and the multipliers are used both for the control group and the treatment groups. B = block; M = multiplier; T = trial.

block, so the multiplier for the first block is included in the index. Again, individual and group differences are readily apparent (Figure 2D) and are readily available for comparison with other relevant individual measures.

Summary and Conclusion

The measures developed in the original article (Gallagher et al., 1993) to better quantify performance in the Morris water maze task included the cumulative search error during learning trials, the average proximity during probe trials, and the spatial learning index to quantify overall learning of the task across days. These measures have had a significant impact in the field and have been instrumental in relating performance on this task with other behavioral measures and neurobiological markers. We propose that a learning index can be used in a variety of other tasks, providing a detailed account of how to construct an index with different task types and different patterns of acquisitions.

References

Barnes, C. A., Suster, M. S., Shen, J., & McNaughton, B. L. (1997). Multistability of cognitive maps in the hippocampus of old rats. *Nature, 388,* 272–275. http://dx.doi.org/10.1038/40859

Baxter, M. G., Bucci, D. J., Gorman, L. K., Wiley, R. G., & Gallagher, M. (1995). Selective immunotoxic lesions of basal forebrain cholinergic cells: Effects on learning and memory in rats. *Behavioral Neuroscience, 109,* 714–722. http://dx.doi.org/10.1037/0735-7044.109.4.714

Beas, B. S., Setlow, B., & Bizon, J. L. (2013). Distinct manifestations of executive dysfunction in aged rats. *Neurobiology of Aging, 34,* 2164–2174. http://dx.doi.org/10.1016/j.neurobiolaging.2013.03.019

Burwell, R. D., & Gallagher, M. (1993). A longitudinal study of reaction time performance in Long-Evans rats. *Neurobiology of Aging, 14,* 57–64. http://dx.doi.org/10.1016/0197-4580(93)90023-5

Burwell, R. D., Saddoris, M. P., Bucci, D. J., & Wiig, K. A. (2004). Corticohippocampal contributions to spatial and contextual learning. *The Journal of Neuroscience, 24,* 3826–3836. http://dx.doi.org/10.1523/JNEUROSCI.0410-04.2004

Castellano, J. F., Fletcher, B. R., Kelley-Bell, B., Kim, D. H., Gallagher, M., & Rapp, P. R. (2012). Age-related memory impairment is associated with disrupted multivariate epigenetic coordination in the hippocampus. *PLoS ONE, 7,* e33249. http://dx.doi.org/10.1371/journal.pone.0033249

Costa, R. M., Federov, N. B., Kogan, J. H., Murphy, G. G., Stern, J., Ohno, M., . . . Silva, A. J. (2002). Mechanism for the learning deficits in a mouse model of neurofibromatosis type 1. *Nature, 415,* 526–530. http://dx.doi.org/10.1038/nature711

Daugherty, A. M., Yuan, P., Dahle, C. L., Bender, A. R., Yang, Y., & Raz, N. (2014). Path complexity in virtual water maze navigation: Differential associations with age, sex, and regional brain volume. *Cerebral Cortex.* Advance online publication. http://dx.doi.org/10.1093/cercor/bhu107

Gallagher, M., Burwell, R., & Burchinal, M. (1993). Severity of spatial learning impairment in aging: Development of a learning index for performance in the Morris water maze. *Behavioral Neuroscience, 107,* 618–626. http://dx.doi.org/10.1037/0735-7044.107.4.618

Goodrich-Hunsaker, N. J., Livingstone, S. A., Skelton, R. W., & Hopkins, R. O. (2010). Spatial deficits in a virtual water maze in amnesic participants with hippocampal damage. *Hippocampus, 20,* 481–491.

Grunwald, I. C., Korte, M., Wolfer, D., Wilkinson, G. A., Unsicker, K., Lipp, H. P., . . . Klein, R. (2001). Kinase-independent requirement of EphB2 receptors in hippocampal synaptic plasticity. *Neuron, 32,* 1027–1040. http://dx.doi.org/10.1016/S0896-6273(01)00550-5

LaSarge, C. L., Montgomery, K. S., Tucker, C., Slaton, G. S., Griffith, W. H., Setlow, B., & Bizon, J. L. (2007). Deficits across multiple cognitive domains in a subset of aged Fischer 344 rats. *Neurobiology of Aging, 28,* 928–936. http://dx.doi.org/10.1016/j.neurobiolaging.2006.04.010

Lee, H. K., Min, S. S., Gallagher, M., & Kirkwood, A. (2005). NMDA receptor-independent long-term depression correlates with successful aging in rats. *Nature Neuroscience, 8,* 1657–1659. http://dx.doi.org/10.1038/nn1586

Maei, H. R., Zaslavsky, K., Teixeira, C. M., & Frankland, P. W. (2009). What is the most sensitive measure of water maze probe test performance? *Frontiers in Integrative Neuroscience, 3,* 4. http://dx.doi.org/10.3389/neuro.07.004.2009

Magnusson, K. R., Scruggs, B., Zhao, X., & Hammersmark, R. (2007). Age-related declines in a two-day reference memory task are associated with changes in NMDA receptor subunits in mice. *BMC Neuroscience, 8,* 43. http://dx.doi.org/10.1186/1471-2202-8-43

Malleret, G., Haditsch, U., Genoux, D., Jones, M. W., Bliss, T. V., Vanhoose, A. M., . . . Mansuy, I. M. (2001). Inducible and reversible enhancement of learning, memory, and long-term potentiation by genetic inhibition of calcineurin. *Cell, 104,* 675–686. http://dx.doi.org/10.1016/S0092-8674(01)00264-1

Miller, S., Yasuda, M., Coats, J. K., Jones, Y., Martone, M. E., & Mayford, M. (2002). Disruption of dendritic translation of CaMKIIalpha impairs stabilization of synaptic plasticity and memory consolidation. *Neuron, 36,* 507–519. http://dx.doi.org/10.1016/S0896-6273(02)00978-9

Morris, R. G. (1981). Spatial localization does not require the presence of local cues. *Learning and Motivation, 12,* 239–260. http://dx.doi.org/10.1016/0023-9690(81)90020-5

Morris, R. G., Garrud, P., Rawlins, J. N., & O'Keefe, J. (1982). Place navigation impaired in rats with hippocampal lesions. *Nature, 297,* 681–683. http://dx.doi.org/10.1038/297681a0

Peleg, S., Sananbenesi, F., Zovoilis, A., Burkhardt, S., Bahari-Javan, S., Agis-Balboa, R. C., . . . Fischer, A. (2010). Altered histone acetylation is associated with age-dependent memory impairment in mice. *Science, 328,* 753–756. http://dx.doi.org/10.1126/science.1186088

Pittenger, C., Huang, Y. Y., Paletzki, R. F., Bourtchouladze, R., Scanlin, H., Vronskaya, S., & Kandel, E. R. (2002). Reversible inhibition of CREB/ATF transcription factors in region CA1 of the dorsal hippocampus disrupts hippocampus-dependent spatial memory. *Neuron, 34,* 447–462. http://dx.doi.org/10.1016/S0896-6273(02)00684-0

Rapp, P. R., & Gallagher, M. (1996). Preserved neuron number in the hippocampus of aged rats with spatial learning deficits. *Proceedings of the National Academy of Sciences of the United States of America, 93,* 9926–9930. http://dx.doi.org/10.1073/pnas.93.18.9926

Robitsek, R. J., Fortin, N. J., Koh, M. T., Gallagher, M., & Eichenbaum, H. (2008). Cognitive aging: A common decline of episodic recollection and spatial memory in rats. *The Journal of Neuroscience, 28,* 8945–8954. http://dx.doi.org/10.1523/JNEUROSCI.1893-08.2008

Rosenzweig, E. S., Rao, G., McNaughton, B. L., & Barnes, C. A. (1997). Role of temporal summation in age-related long-term potentiation-induction deficits. *Hippocampus, 7,* 549–558. http://dx.doi.org/10.1002/(SICI)1098-1063(1997)7:5<549::AID-HIPO10>3.0.CO;2-0

Sandstrom, N. J., Kaufman, J., & Huettel, S. A. (1998). Males and females use different distal cues in a virtual environment navigation task. *Cognitive Brain Research, 6,* 351–360. http://dx.doi.org/10.1016/S0926-6410(98)00002-0

Schoenbaum, G., Nugent, S., Saddoris, M. P., & Gallagher, M. (2002). Teaching old rats new tricks: Age-related impairments in olfactory reversal learning. *Neurobiology of Aging, 23,* 555–564. http://dx.doi.org/10.1016/S0197-4580(01)00343-8

Smith, T. D., Adams, M. M., Gallagher, M., Morrison, J. H., & Rapp, P. R. (2000). Circuit-specific alterations in hippocampal synaptophysin immunoreactivity predict spatial learning impairment in aged rats. *The Journal of Neuroscience, 20,* 6587–6593.

Spiegel, A. M., Koh, M. T., Vogt, N. M., Rapp, P. R., & Gallagher, M. (2013). Hilar interneuron vulnerability distinguishes aged rats with

memory impairment. *The Journal of Comparative Neurology, 521,* 3508–3523. http://dx.doi.org/10.1002/cne.23367

Tomás Pereira, I., Coletta, C. E., Perez, E. V., Kim, D. H., Gallagher, M., Goldberg, I. G., & Rapp, P. R. (2013). CREB-binding protein levels in the rat hippocampus fail to predict chronological or cognitive aging. *Neurobiology of Aging, 34,* 832–844. http://dx.doi.org/10.1016/j.neurobiolaging.2012.07.010

Wilson, I. A., Ikonen, S., McMahan, R. W., Gallagher, M., Eichenbaum, H., & Tanila, H. (2003). Place cell rigidity correlates with impaired spatial learning in aged rats. *Neurobiology of Aging, 24,* 297–305. http://dx.doi.org/10.1016/S0197-4580(02)00080-5

Yang, S., Megill, A., Ardiles, A. O., Ransom, S., Tran, T., Koh, M. T., . . . Kirkwood, A. (2013). Integrity of mGluR-LTD in the associative/commissural inputs to CA3 correlates with successful aging in rats. *The Journal of Neuroscience, 33,* 12670–12678. http://dx.doi.org/10.1523/JNEUROSCI.1086-13.2013

Received January 31, 2015
Revision received May 6, 2015
Accepted May 23, 2015 ■