



# A tale of too many tasks: task fragmentation in motor learning and a call for model task paradigms

Rajiv Ranganathan<sup>1</sup> · Aimee D. Tomlinson<sup>1</sup> · Rakshith Lokesh<sup>1</sup> · Tzu-Hsiang Lin<sup>1</sup> · Priya Patel<sup>1</sup>

Received: 31 January 2020 / Accepted: 17 August 2020 / Published online: 10 November 2020  
© Springer-Verlag GmbH Germany, part of Springer Nature 2020

## Abstract

Motor learning encompasses a broad set of phenomena that requires a diverse set of experimental paradigms. However, excessive variation in tasks across studies creates fragmentation that can adversely affect the collective advancement of knowledge. Here, we show that motor learning studies tend toward extreme fragmentation in the choice of tasks, with almost no overlap between task paradigms across studies. We argue that this extreme level of task fragmentation poses serious theoretical and methodological barriers to advancing the field. To address these barriers, we propose the need for developing common ‘model’ task paradigms which could be widely used across labs. Combined with the open sharing of methods and data, we suggest that these model task paradigms could be an important step in increasing the robustness of the motor learning literature and facilitate the cumulative process of science.

**Keywords** Replication · Power · Design · Standardization · Sample size · Science

## Introduction

In the context of theories in psychology, Mischel referred to the presence of a ‘toothbrush problem’—i.e., that “no self-respecting person wants to use anyone else’s” (Mischel 2008). In this review, we show that experimental paradigms in motor learning suffer from a similar toothbrush problem; tasks used are extremely fragmented, with little to no overlap between different studies. We argue that this extreme fragmentation has negative consequences for the field from both theoretical and methodological perspectives. Finally, we propose the use of common ‘model task paradigms’ to address these issues.

## How fragmented is motor learning?

### Review of motor learning studies

To quantify the degree of fragmentation in the motor learning literature, we selected publications in 2017 and 2018 from the following five journals: *Experimental Brain Research*, *Human Movement Science*, *Journal of Motor Behavior*, *Journal of Motor Learning and Development* and *Journal of Neurophysiology*. We included a study in our analysis if (i) there was change in motor performance arising due to practice, (ii) there were at least two a priori defined groups which were compared in a between-subject design on the same task, and (iii) the population included was unimpaired adults. Overall, 64 papers were included based on these criteria.

These criteria were based on the following rationale. Our focus was on motor learning studies from a behavioral emphasis, which led to us select journals that publish regularly on this theme. The timespan of 2 years was to provide a reasonable sample of studies (our goal was to do at least 50). Because one of our particular emphasis was to examine issues related to design and sample size, the two-group minimum requirement was used so that our analysis could focus on the more common ‘between-group’ designs (rather than ‘within-subject’ designs). Also, studies with within-subject

---

Communicated by Patrick Haggard.

✉ Rajiv Ranganathan  
rrangana@msu.edu

<sup>1</sup> Department of Kinesiology, Michigan State University, 308 W Circle Dr, East Lansing, MI 48824, USA

designs in motor learning tend to typically describe changes associated with learning (e.g. changes in EMG or the use of redundancy) rather than compare different practice strategies, which make them less relevant to the central arguments related to fragmentation being advanced here. Finally, because tasks often have to be modified for children or adults with movement impairments, we focused on studies with unimpaired adults to get an estimate of fragmentation in cases where the tasks do not necessarily have to be modified.

For each study, we examined the task used and classified it into one of six categories—adaptation, applied, coordination, sequence, tracking, and variability. This classification was based on prior work that have highlighted such distinctions (e.g. scaling vs. coordination (Newell 1991), or adaptation vs. skill learning (Krakauer and Mazzoni 2011)), but these prior distinctions have typically only been dichotomous. So, we expanded these into six categories to more accurately group the types of tasks in the studies that were reviewed. The broad rules for each of these categories are specified in Table 1. Five of these categories were based on the types of learning involved in the task (adaptation, coordination, sequence, tracking and variability), while the ‘applied’ category was used for studies where either the task itself was central to the research question or it was used without modification from its implementation in real-world settings. For example, a golf putting task in a lab was categorized under ‘variability’, but a study examining professional golf players putting on an actual green was categorized as ‘applied’. We adopted this category because we felt that these studies, where the task is integral to the research question, are less likely to benefit from the development of a common task paradigm. Moreover, although these categories were not always mutually exclusive, for the sake of

this review, we classified each study into only one category. When there were conflicts on the categorization, the authors resolved such conflicts through discussion.

We then examined the actual tasks used in each category to determine if these studies used the same experimental paradigm. For each task, we extracted relevant parameters that were reported in the Methods section regarding the implementation of the task. For example, when reaching was used as a task in the ‘adaptation category’, we compared experimental parameters such as the type of perturbation, the amplitude of the reach, the number of targets, and instructions to participant such as whether they had to stop inside the target or shoot through it) Similarly, for a sequence learning task we examined the number of elements in the sequence, how many fingers were used, the instructions to the participant such as “do a fixed number of sequences” vs. “do as many sequences in a fixed time” etc.). It is important to note that these coded features were only based on the task itself, and did not include differences in protocol information (such as the amount or duration of practice). Based on this information, a task paradigm was classified as ‘unique’ if it did not match any other task paradigm in its category. In addition to the task paradigm, the sample size/group was also calculated. If there were multiple experiments, the sample size/group was simply computed as the total sample size/total number of groups. This information for each task along with the coded features is summarized in Table 2.

Overall, we found that a majority of the studies on motor learning were on adaptation (36%), followed by coordination (17%), sequence learning (16%), variability (12%), tracking (11%), and applied tasks (8%) (Fig. 1a). Group sample sizes were typically in the range of 10–16 participants (25th–75th percentile) (Fig. 1b). Critically, we

**Table 1** Definition of task categories

Task category	Process involved in learning of task	Typical dependent variable(s)	Example
Adaptation	Responding to perturbations of typically well-learned movements	Deviation from baseline behavior	Visuomotor rotation
Applied	Production of movement responses in ‘real-world’ situations that may involve a combination of processes	Task-dependent	Penalty kicks in soccer
Coordination	Production of spatiotemporal pattern involving more than a single degree of freedom (limbs, joints, muscles)	Coordination measures that capture relative motion between the degrees of freedom, dimensionality reduction techniques	Bimanual coordination
Sequence	Production of a sequence of several movement responses	Speed, errors, reaction time	Key pressing
Tracking	Production of a desired spatiotemporal pattern that is ‘time varying’	Deviation between the target pattern and the actual	Tracking a moving target on a screen
Variability	Production of a ‘steady-state’ task performance level over time or trials	Variability across time or trials	Throwing a ball to a target

Tasks used in prior studies were classified into one of six categories based on the processes involved during motor learning. Typical dependent variables and a prototypical example of such tasks is also provided. Even though categories themselves are not completely mutually exclusive, each study in the review was classified into one of the six categories

**Table 2** Review of motor learning studies from selected journals in 2017–2018

Authors	Category	Task	Task notes	Sample Size Group	Sample Size per Group	Unique Task (Y/N)
Couth et al. (2018)	Adaptation	Grasping	Grasping haptic blocks of different lengths with vision-haptic conflict	48 (2 groups)	24	Y
Megregor et al. (2018)	Adaptation	Reaching	FF curl k = 14 Ns/m. Amplitude unspecified, Timing 375 ± 100 ms, 1 target	78 (3 groups)	26	Y
McGregor and Gribble (2017)	Adaptation	Reaching	FF curl k = 14 Ns/m. Amplitude 15 cm, Timing 450–550 ms, 1 target	30 (2 groups)	15	Y
McKenna et al. (2017)	Adaptation	Reaching	FF curl k = 15 Ns/m. Amplitude 10 cm, Peak velocity between 0.25–0.35 m/s, time < 750 ms, 1 target	84 (6 groups)	14	Y
Milner et al. (2018)	Adaptation	Reaching	FF curl k = 15 Ns/m. Amplitude 25 cm, Timing 600 ± 50 ms, 2 targets	28 (2 groups)	14	Y
Melendez-Calderon et al. (2017)	Adaptation	Reaching	FF curl k = 25 Ns/m. Amplitude 10 cm, Timing 500–700 ms, 3 targets	Exp1: 7 (1 group); Exp2: 30 (3 groups)	9.25	Y
Kumar et al. (2018)	Adaptation	Reaching	VMR 30 deg CW and CCW. Amplitude 13 cm, 8 targets (45 deg apart)	48 (6 groups)	8	Y
Schmitz and Bock (2017)	Adaptation	Reaching	VMR 30 deg CW and CCW; Amplitude 36 cm, 6 visual and auditory targets (+-6/18/30 degs)	36 (3 groups)	12	Y
Aiken et al. (2017)	Adaptation	Reaching	VMR 30 deg CW, Amplitude 3 cm, 3 targets (60, 150 and 240 deg)	49 (4 groups)	12.25	Y
Holland et al. (2018)	Adaptation	Reaching/Shooting	VMR 15 and 25 deg CW and CCW; Amplitude 10 cm, 200–1000 ms time, 1 target (90 deg)	Exp1: 40 (3 groups); Exp 2: 20 (2 groups)	12	Y
Canaveral et al. (2017)	Adaptation	Reaching/Shooting	VMR 15 deg CCW; Amplitude 10 cm, 150 ms time, 8 targets (45 degs starting from 22.5)	Exp1: 32 (3 groups); Exp 2: 23 (2 groups)	11	Y
Jiang et al. (2018)	Adaptation	Reaching/Shooting	VMR 30 deg CCW, Amplitude 7 cm, timing 100–400 ms, 1 target	192 (16 groups)	12	Y
Jalali et al. (2017)	Adaptation	Reaching/Shooting	VMR 30 deg CCW, Amplitude 8 cm, timing of 100–300 ms, 8 targets	Exp1: 28 (2 groups); Exp 2: 20 (2 groups); Exp 3: 27(2 groups); Exp 4: 24 (2 groups); Exp 5: 36 (2 groups); Exp 6: 32 (2 groups); Exp7: 26 (2 groups)	13.79	Y
Leow et al. (2017)	Adaptation	Reaching/Shooting	VMR 30 deg CW and CCW, Amplitude 9 cm, 8 targets	Exp 1A: 14 (1 group); Exp 1B: 14 (1 group); Exp 2: 36 (3 groups); Exp 3: 10 (1 group)	12.33	Y
Song and Smiley-Oyen (2017)	Adaptation	Reaching/Shooting	VMR 50 deg CW, Amplitude 10 cm, Timing 100–600 ms, 1 target	44 (4 groups)	11	Y

Table 2 (continued)

Authors	Category	Task	Task notes	Sample Size	Sample Size per Group	Unique Task (Y/N)
Reilly and Pettibone (2017)	Adaptation	Speech	5 CVC (consonant–vowel–consonant) words with different embedding of vowels	40 (2 groups)	20	Y
Seidler et al. (2016)	Adaptation	Throwing	Prism goggles—40D, dart throw from 7ft	60 (4 groups)	15	Y
Cattan et al. (2018)	Adaptation	Tracking	Delay adaptation, Tracing ellipse and a Fig. 8 pattern	20 (2 groups) in Exp 1; 26 (2 groups) in Exp 2	11.5	Y
French et al. (2018)	Adaptation	Walking	Distorted visual feedback of stride length by upto 9%, Treadmill Speed—1.2 m/s	Exp 1: 28 (2 groups); Exp 2: 6 (1 group)	11.33	Y
Maeda et al. (2016)	Adaptation	Walking	Prism goggles—20D, Precision walking task across 6-m path	100 (10 groups)	10	Y
Hinkel-Lippsker and Hahn (2017)	Adaptation	Walking	Split-belt treadmill with one limb $\pm 0.5$ m/s of self-selected walking speed	48 (6 groups)	8	Y
Hinkel-Lippsker and Hahn (2018)	Adaptation	Walking	Split-belt treadmill with one limb $\pm 0.5$ m/s of self-selected walking speed	48 (3 groups)	16	N
Day et al. (2018)	Adaptation	Walking	Split-belt treadmill with one limb at 0.7, one limb at 1.4 m/s	40 (4 groups)	10	Y
Kim et al. (2018)	Applied	Golf putting	2 m, 5 m putts on real green	20 (2 groups)	10	Y
Navarro et al. (2018)	Applied	Penalty kicks		20 (2 groups)	10	Y
Chien and Chen (2018)	Applied	Pétanque		75 (2 groups)	37.5	Y
Krajenbrink et al. (2018)	Applied	Slingerbball	5 m throw to a 2 m dia target	162 (2 groups)	81	Y
Hebert (2018)	Applied	Speed cup stacking		117 (3 groups)	39	Y
Lin et al. (2018)	Coordination	Arm-Trunk reaching	Reaching for targets with trunk motion	48 (4 groups)	12	Y
Panzer et al. (2018)	Coordination	Bimanual	1:2 pattern using forearm	32 (4 groups)	8	Y
Bingham et al. (2018)	Coordination	Bimanual	90 deg rel phase using joysticks	20 (2 groups)	10	Y
Dyer et al. (2017)	Coordination	Bimanual	Bimanual shape tracking with music feedback	60 (3 groups)	20	Y
Cardis et al. (2017)	Coordination	Bimanual	Throwing virtual puck with both hands toward a target	50 (5 groups)	10	Y
Thorpe et al. (2016)	Coordination	Bimanual	Virtual reaching task with IMUs	16 (2 groups)	8	Y
Kimura et al. (2017)	Coordination	Dual task of isometric knee extension and elbow flexion	Knee extension required force production; Reaction time task with elbow	45 (3 groups)	15	Y
Shuggi et al. (2018)	Coordination	Head-control reaching	Reaching to targets with a robot arm while avoiding boundaries	20 (2 groups)	10	Y

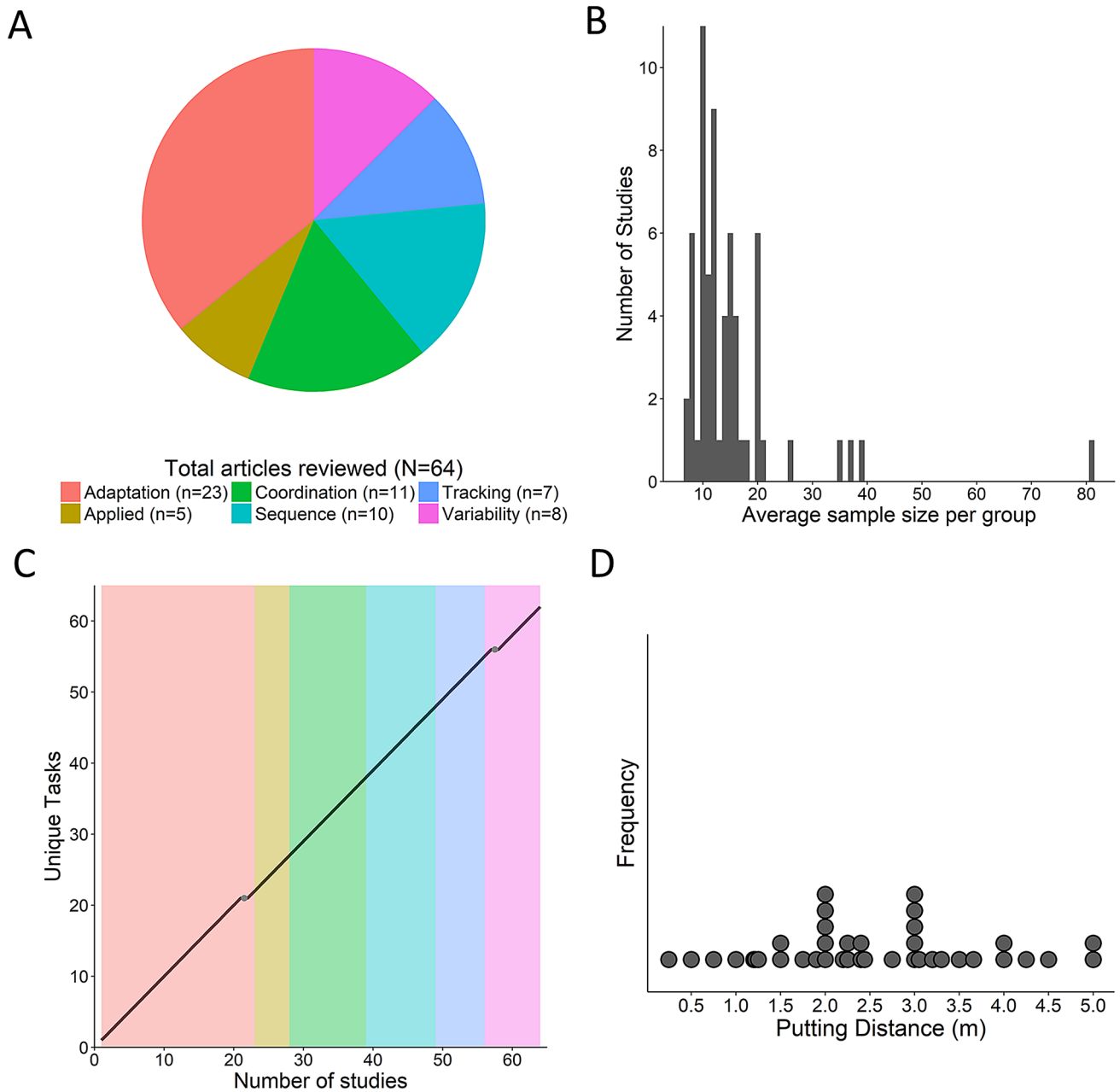
**Table 2** (continued)

Authors	Category	Task	Task notes	Sample Size Group	Sample Size per Group	Unique Task (Y/N)
Chua et al. (2018)	Coordination	Maximum vertical jump	Height measured using a wearable accelerometer	36 (2 groups)	18	Y
LoJacono et al. (2018)	Coordination	Obstacle crossing	Real and virtual environments	40 (2 groups)	20	Y
Wulf et al. (2018)	Coordination	Throwing	Lasso to a target 10 feet away	Exp1: 32 (2 groups) Exp2: 42 (3 groups)	14.8	Y
Krause et al. (2018)	Sequence	Elbow flexion–extension	70, 20, 70 degs in 1200 ms	42 (2 groups)	21	Y
Sobierajewicz et al. (2017)	Sequence	Keypress	5 element keypress sequence with both single and 4 fingers- as fast as possible but for a fixed number of sequences	24 (2 groups)	12	Y
Neville and Trempe (2017)	Sequence	Keypress	5 element keypress sequence with 4 fingers—as fast as possible without pause in 30 s	42 (3 groups)	14	Y
Yokoi et al. (2016)	Sequence	Keypress	5 element keypress sequence with 5 fingers- as fast without pause for fixed number of sequences	32 (4 groups)	8	Y
Karlinsky and Hodges (2018)	Sequence	Keypress	5 element keypress sequence with one finger and specified movement times	94 (6 subgroups)	15.67	Y
Ruffino et al. (2017)	Sequence	Peg placing	Modified version of the Nine Hole Peg Test	20 (2 groups)	10	Y
Barros et al. (2017)	Sequence	Reaching	3 item sequence with specified times (Exp1) or specified forces (Exp2)	Exp1: 60 (4 groups); Exp2: 60 (4 groups)	15	Y
Kaipa et al. (2017)	Sequence	Speech	Non meaningful phrase with time constraints	80 (4 groups)	20	Y
Levac et al. (2017)	Sequence	Stepping	6 step sequences	40 (2 groups)	20	Y
Haarman et al. (2017)	Sequence	Walking	7 step sequence (6 step lengths) with targets projected on to treadmill	20 (2 groups)	10	Y
Dotov and Froese (2017)	Tracking	Auditory-motor synchronization task	Hand-held accelerometer movements created sounds that had to be matched to target sound	48 (3 groups)	16	Y
Itaguchi and Fukuzawa (2018)	Tracking	Drawing	Draw a specific curved shape in one stroke	14(2 groups)	7	Y
Fialho and Tresilian (2018)	Tracking	Interception	Hitting a moving target using a manipulandum	21 (3 groups)	7	Y
Yen et al. (2018)	Tracking	Reaching	Move cursor in a virtual maze using isometric force	20 (2 groups)	10	Y
Crespo et al. (2017)	Tracking	Tracking	Tracking a virtual loop through a continuous or discrete wire	30 (3 groups)	10	Y

Table 2 (continued)

Authors	Category	Task	Task notes	Sample Size	Sample Size per Group	Unique Task (Y/N)
Boyer et al. (2017)	Tracking	Tracking	Tracking randomly moving target with sonification of errors	Exp 2: 36 participants (3 groups); Exp 3: 36 participants (3 groups); 12 participants (1 control group)	12	Y
Kaipa and Kaipa (2018)	Tracking	Tracking	Used electromyogram from mouth to track ramp of upto 40% maximum voluntary contraction	45 (3 groups)	15	Y
Mansfield et al. (2017)	Variability	Balance	Oscillating platforms with varying frequencies	48 (3 groups); 15/15/18	16	Y
Ong and Hodges (2018)	Variability	Balance	Stand on Stabilometer	Exp1: 48 (4 groups); Exp2: 40 (4 groups)	11	Y
Meira and Fairbrother (2018)	Variability	Balance	Stand on Stabilometer	Exp1: 56 (4 groups); Exp2: 48 (4 groups)	13	N
Frank et al. (2018)	Variability	Golf putting	3 m putt to hole on artificial green	24 (2 groups)	12	Y
Raisbeck and Diekfuss (2017)	Variability	Pistol shooting	15 feet from target; 2.5/5 cm target	68 (4 groups)	17	Y
Pacheco and Newell (2018)	Variability	Throwing	Ball toss to hexagonal and square targets placed 2.05 m away	24 (3 groups)	8	Y
Gimneken et al. (2018)	Variability	Throwing	Ball toss to target 4 m away	40 (2 groups)	20	Y
Grand et al. (2017)	Variability	Throwing	Beanbag toss over a board to target 3 m away	70 (2 groups)	35	Y

FF Force field, VMR visumotor rotation, CW clockwise, CCW counterclockwise



**Fig. 1** Fragmentation of task paradigms in motor learning. **a** Distribution of task categories across all studies—studies were divided into six categories based on the criteria shown in Table 1. **b** Distribution of sample size/group across all studies. The majority of studies ranged between 10–16 participants in a group. **c** Plot of the number of studies reviewed in the current paper on the x-axis against the number of uniquely different tasks on the y-axis. The two dots high-

light the two studies that were not classified as unique. The almost perfect diagonal line indicates that there was little to no overlap in tasks between studies across all six categories, **(d)** Variability in putt distance across golf putting tasks shows fragmentation of task paradigms even within the same task. The peak at 3 m is almost exclusively driven by a single author group

found that out of the 64 studies reviewed, there were 62 unique task paradigms, a fragmentation of ~97% (Fig. 1c). In fact, of the two matches found, one was from the same group of authors, and the second was a standard commercially available device. Although many of studies used task paradigms that belong to what could be considered

the same task category (e.g., visuomotor adaptation in reaching), they varied in the implementation of the task in terms of relevant task parameters (see Table 2). These results highlight the high degree of fragmentation in task paradigms in the field.



## Review of task fragmentation within the same task—golf putting

One possible reason for the high level of task fragmentation could be a consequence of the fact that motor learning is a diverse field, and tasks are tailored for specific research questions. Therefore to examine if task fragmentation exists even when researchers seemingly choose the *same* task, we performed a second analysis where we focused on studies that all used a relatively common task—golf putting. We used the search phrases ‘golf’ and ‘motor learning’ in Web of Science, and studies from 2013–2018 were examined (we changed the timeframe to increase the number of articles that were included). The same inclusion criteria were used as before, and we only focused on putting studies (e.g. studies on the golf swing or chip shots were not included). The relevant parameters that were extracted for this task were the putt distance, target type, target size, the surface type, and the scoring system were extracted (Table 3). The scoring system, while not strictly part of the task itself, was included as a factor because it has a direct influence on how results from different studies can be interpreted relative to each other. Overall 22 studies were selected.

Once again, results showed that even within the context of the same task, studies used a variety of different putt distances (Fig. 1d). Even though there were several studies that used a putting distance of 3 m, this was almost exclusively driven by a single group of authors. In addition, there were also variations in the hole type, diameter, and the scoring system (Table 3). As emphasized earlier, even though such differences may seem trivial at first glance, they can create important differences between studies. For example, the use of an actual hole (as opposed to a circle) increases ecological validity but adversely affects estimation of error and variability. This is because a range of ball velocities will land the ball in the hole and be ‘compressed’ as zero error. Similarly, from a scoring standpoint discretizing a continuous error measure, such as the distance of the ball from the center of the target into measures such as the number of successful putts or a points system, has the potential to significantly distort learning curves (Bahrck et al. 1957). Finally, other information was incomplete to the extent that it would be difficult to perform a direct replication. For example, in the putting surface, the term ‘synthetic’ or ‘artificial’ green was used in several papers; however, only two papers mentioned the ‘speed’ of the green used, which is a critical variable for replication. These results highlight that even when the same general task is chosen, there is still a high degree of fragmentation in task paradigms across studies.

## Consequences of task fragmentation

Given the evidence for fragmentation, one argument could be that these variations in parameters are trivial and do not alter our understanding of motor learning in any meaningful way. We wish to highlight two things in this regard—(i) there is evidence that at least some of these trivial parameters can have a significant influence on the dependent variables, which in turn may affect inferences about learning. For example, in visuomotor rotation, the choice of the number of targets and the rotation angle has been shown to influence the rate of learning and the magnitudes of implicit and explicit learning (Bond and Taylor 2015) (ii) in rare cases, these parameter changes do have the potential to completely alter the conclusions observed. For example, the presence or absence of catch trials during training in force field adaptation has been shown to influence whether motor memory consolidation is observed (Overduin et al. 2006). Given these possibilities, we describe both the theoretical and methodological consequences of task fragmentation in detail.

### Theoretical consequences

From a theoretical standpoint, the fragmentation of tasks across studies makes every finding an ‘island’ and hampers the cumulative progress of science, where researchers can replicate the findings and then build off of previous research (Zwaan et al. 2018). This issue is critical in light of the recent ‘replication crisis’, where well-known effects in many fields have either failed to replicate or have had smaller effects than once originally assumed (Camerer et al. 2018; Open Science Collaboration 2015). Although one argument for using different tasks is that they can be useful in testing the generality of theories and hypotheses in different contexts, the utility of *exclusively* relying on such ‘conceptual replications’ has been challenged because it is subject to confirmation bias (Chambers 2017). In other words, if the results of the original and the conceptual replication agree, then it is taken as evidence of the generality of the effect; however, if the results of the original study and the conceptual replication differ (i.e. the replication ‘failed’), then these differences are directly attributed to the changed task parameters in the conceptual replication. This leads to a situation where the robustness of the original finding can never be questioned. Moreover, it has been argued that even when such conceptual replications fail, current incentive structures at many journals make the likelihood of publishing these results low, leading to a biased literature (Pashler and Harris 2012).

This issue of being able to exactly replicate a study (i.e. a direct replication) is especially important in the context



**Table 3** Review of selected motor learning studies using golf putting as a task in 2013–2018

Authors	Putting distance	Target type	Target size	Description of surface	Performance measure(s)
Chauvel et al. (2013)	0.25, 0.5, 0.75, 1, 1.25, 1.5, 1.75, 2, 2.25	Hole	11.4 cm dia	Artificial turf indoor green 2 m × 2.7 m	Number of successful putts
Daou et al. (2016)	1.2 m	Point	target cross made up of two 10.8 cm pieces of tape	Artificial grass surface	Radial error, bivariate error
Fazeli et al. (2017)	1.22, 2.44, 3.66 m	Circle	4 cm dia	Synthetic 9 m × 4 m putting green (STI Pro Putt)	Radial error measured from edge of target
Palmer et al. (2016)	1.5 m	Square	2 × 2 cm square	4 m × 0.55 m level artificial-turf indoor green	Distance between center of target and edge of ball; distance counted as 100 cm if rear border of green was contacted
Zhu et al. (2015)	1.9 m	Hole	12 cm dia	Artificial grass putting surface	Number of successful putts
Chauvel et al. (2015)	2 m	Circle	10.4 cm dia	4 m × 0.55 m level artificial-turf indoor green	Distance between center of target and edge of ball; distance counted as 100 cm if rear border of green was contacted
Ziv and Lidor (2015)	2 m	Circle	10.8 cm dia	1 m × 4 m artificial grass	Absolute error, variable error
Abbas and North (2018)	2, 5 m; target behind opaque curtain	Hole	10.61 cm dia	Flat carpeted surface	Radial error
Land et al. (2014)	2.0, 2.75, 3.5, 4.25, 5.0 m	Circle	10.8 cm dia	4 m × 9 m artificial indoor putting green	Radial error, Bivariate radial error
Nunes et al. (2019)	2.2, 2.4, 3.2 m	Hole	N/A	1.5 m × 4 m artificial grass	Point system
Lawrence et al. (2014)	2.25 m	Hole	10 cm dia	Astroturf surface, 90-cm incline slope of 22 degrees that started 72 cm from the start position	Number of successful putts, radial error
Ring et al. (2015)	2.4 m	Hole	10.8 cm dia	Flat artificial putting surface (Turf tiles) with stimpmeter reading of 3.05 m	Number of putts, radial error; zero radial error if putt was holed
Daou et al. (2019)	3 m	Point	target cross made up of two 10.8 cm pieces of tape	Artificial grass surface	Radial error, bivariate variable error
Frank et al. (2013)	3 m	Circle	10.8 cm dia	4 m × 7 m artificial indoor putting green	Radial error, bivariate variable error
Frank et al. (2014)	3 m	Circle	10.8 cm dia	Artificial indoor putting green	Radial error, bias, bivariate variable error
Frank et al. (2016)	3 m	Circle	10.8 cm dia	4 m × 9 m artificial indoor putting green	Radial error, bivariate variable error
Kim et al. (2017)	3 m	Circle	10.8 cm dia	4 m × 9 m synthetic putting green	Radial error
Vine et al. (2013)	3.05 m	Hole	10.8 cm dia	6 × 2.5 m artificial putting green, stimpmeter reading 3.28 m	Radial error; zero if putt was holed
Krause et al. (2014)	3.3 m; putt from behind a curtain	Point	N/A	Carpet with comparable conditions to a green	Absolute error in distance and angle
Kearney (2015)	4 m	Circle	10 cm dia	Indoor putting surface	Points system
Lewthwaite et al. (2015)	4 m	Circle	20 cm dia	Floor	Points system
Munzert et al. (2014)	4.50 m	Point	N/A	Indoor green	Absolute error

of motor learning because of the recognition of the role of tasks and task constraints in determining behavior (Newell 1986, 1989). Given that tasks can vary along several dimensions, it is perhaps not surprising that many of these dimensions have been used as explanations for discrepancies in results across studies—e.g., practice spacing effects are affected by whether tasks are discrete or continuous (Lee and Genovese 1989), frequency of augmented feedback effects are affected by whether tasks are simple or complex (Wulf and Shea 2002), and sleep consolidation effects are affected by whether the tasks involve sequence learning or adaptation (Doyon et al. 2009). Although these task dimensions certainly play an important role, it is also important to recognize that the true effect that they play cannot be fully understood without ensuring the robustness of the original findings through direct replications.

### Methodological consequences

In addition to theoretical consequences, there are also methodological consequences of task fragmentation. Here, we focus on three primary consequences of such fragmentation—(i) arbitrariness in choice of task parameters, (ii) arbitrariness in choice of sample size, and (iii) inability to compare magnitudes of effects across different manipulations.

#### Arbitrariness in task parameters.

At the experimental design stage, the use of a new task poses a challenge for the experimenter because it requires making choices about several task parameters that are critical to the experiment without sufficient information. For example, in a motor learning study, a common parameter that is a critical part of the experiment is the practice duration; yet this choice is rarely explicitly justified. Instead, researchers are likely to choose these values through a combination of unpublished pilot testing, applying rules of thumb based on other published studies, and convenience (e.g., choosing the shortest duration possible). These arbitrary choices can greatly limit the generalization of motor learning findings to the real-world—for example, in spite of the seemingly diverse range of tasks used, very few studies focus on the period after extended practice, when an effective performance plateau has been reached (Hasson et al. 2016).

An even bigger challenge is the choice of the experimental manipulation itself. Typically, studies in motor learning involve a between-group manipulation of an independent variable with the experimenter having to decide on what the values of these variable are. For example, a study on variable practice may use a throwing task and compare two groups—a ‘constant’ group that always practices throws to a target from the same distance, and a ‘variable’ group that practices throws to different distances (Kerr and Booth

1978). However, the critical parameter choice of how much variability the variable group experiences can have a major influence on the observed results (Cardis et al. 2018). This is because most practice strategies (e.g., practice variability, practice spacing, feedback frequency) are likely to be ‘non-monotonic’ with respect to their influence on learning—i.e., there is an optimal level that maximizes learning, and there is a decrease in the amount of learning both above and below this optimal level (Guadagnoli and Lee 2004).

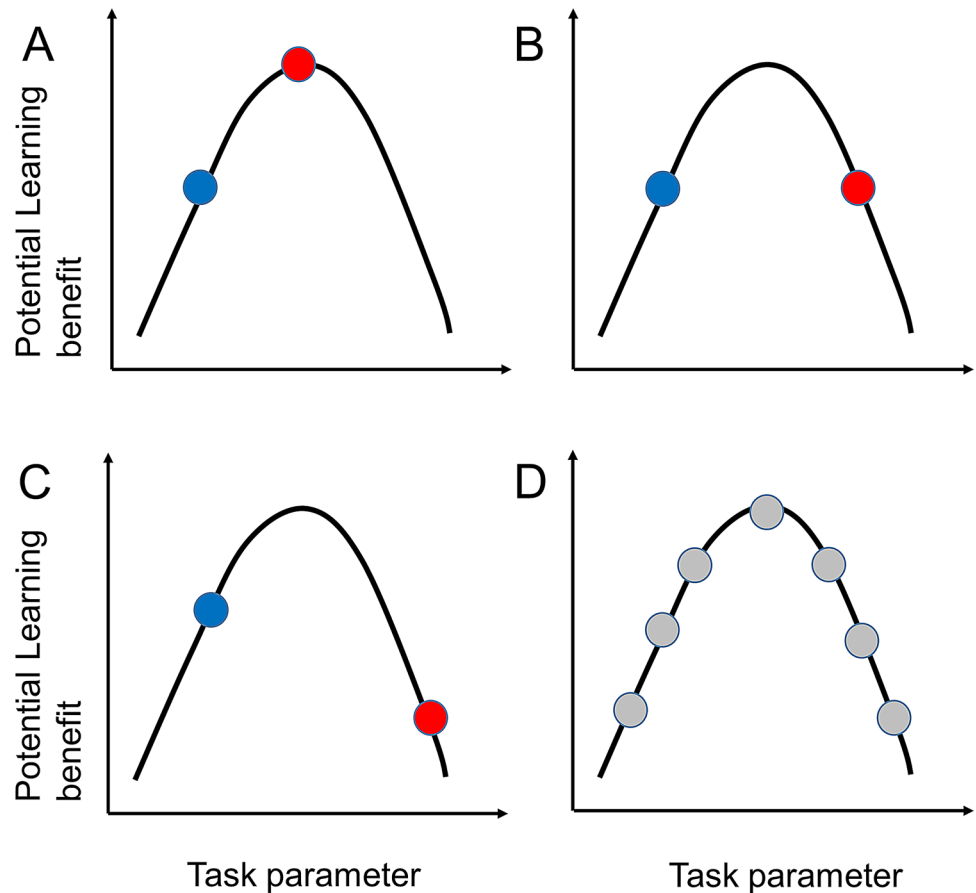
When the choice of this task parameter is made without sufficient information, it means that the experimenter does not know where the groups lie on this non-monotonic function (Fig. 2). As a result, even when the underlying effect of a manipulation is consistent across studies, different studies may get seemingly ‘contradictory’ results simply because they are sampling at different points of this function (Fig. 2a–c). To further compound this problem, when studies use different tasks, these discrepancies in results caused by variations in sampling may get incorrectly attributed to the task itself. One potential solution for this problem is to characterize a full ‘dose–response curve’ for each task and task manipulation by increasing the number of groups and sampling across the full parameter range (Fig. 2d). However, given the amount of effort needed to establish this dose–response curve, doing it for every new task would likely be impractical. These highlight the need for fewer tasks, and more data on these tasks to make more informed decisions on these task parameters.

#### Arbitrariness in sample size planning

Another key parameter choice in experimental design is the sample size. Several reviews have emphasized the need for a priori sample size planning, because the lack of adequate power stemming from small sample sizes can greatly reduce the reliability of the findings in the literature (Button et al. 2013). However, just like the task parameters, sample size planning tends to become arbitrary when a new task is used. Consistent with a prior review of studies in motor learning (Lohse et al. 2016), sample sizes seen in our current review were around the 10–16/group (25th–75th percentile), regardless of the effect being studied. These sample sizes suggest that they are likely driven by heuristics for meeting a ‘publication threshold’ for journals.

The major barrier to doing a priori sample size planning in a new task is the lack of information on the expected effect size, or alternatively the ‘smallest effect size of interest’ (Lakens et al. 2018). Effect sizes estimated from small samples of pilot data are not reliable (Albers and Lakens 2018), and even meta-analytic estimates of effect size in motor learning seem to suffer from issues of small sample size and publication bias (Lohse et al. 2016). Moreover, as mentioned in the task parameters section, heterogeneity in

**Fig. 2** Arbitrary selection of task parameters can lead to seemingly ‘contradictory’ results. Considering that the effect of many manipulations in motor learning (e.g., variability, spacing, feedback frequency etc.) are likely non-monotonic, task parameter selection becomes critical. For example, in a simple two group design (the ‘low’ group in blue, and the ‘high’ group in red), even though the underlying relation is the same in all cases, arbitrary selection of the task parameter can lead to (a) the high group learning more than the low group, (b) no difference between groups, or (c) the low group learning more than the high group. **d** A dose–response curve with more groups avoids this issue by providing a complete description of the non-monotonic relation, however given the much larger sample size required, establishing such dose–response curves becomes impractical when each study uses a different task



the effect size across different studies could also be driven by factors such as variation in the tasks and task parameters. These issues again point to the need for more data on fewer tasks to obtain reliable estimates of effect sizes, and thereby determine an appropriate sample size.

**Uninformative magnitude of effects**

Finally, if motor learning studies are to be of relevance to the real world (and not just mere demonstrations of effects), the goal is not only to detect ‘if’ there is an effect of an intervention, but also to estimate the size of the effect—i.e., what can cause the ‘biggest’ effect. A literature with fragmented tasks is detrimental to this goal because it prevents any relative comparisons of magnitude of effects across different experimental manipulations. For example, how do we compare the relative benefits of manipulations like practice spacing, variable practice, or self-controlled practice if each of these manipulations uses a different task with different dependent variables? Although this might seem like an ‘apples to oranges’ comparison, knowing at least the effective range of performance improvements for each of these manipulations is critical to determining an effective training paradigm in the real world. For example, in rehabilitation studies, this

comparison between different types of therapies (e.g., robotics vs. conventional therapy) is routinely done by comparing the improvement in movement impairment measured on a common scale (e.g. the Fugl-Meyer score) (Lo et al. 2010). However, in motor learning studies, this common scale requires the use of the same task because (i) unlike measures of movement impairment, measures of motor learning, by definition, are specific to the task, and (ii) using standardized effect sizes (e.g. Cohen’s d) to make comparisons across tasks can be problematic because factors other than the mean difference, such as sample variability influence these effect sizes (Baguley 2009). These issues highlight that for comparing magnitudes of effects between different manipulations, there is a need for common tasks (and associated dependent variables), where improvements in performance can be directly compared across studies in raw units of measurement (e.g., error measured in centimeters).

**A call for ‘model’ task paradigms**

In light of these issues, we feel there is a necessity to create a few ‘model’ task paradigms for motor learning studies. These model task paradigms could serve as a common

testbed for several studies on motor learning and labs all over the world could run identical (or nearly identical) experiments using these paradigms. This proposal is analogous to the study of model organisms in related fields like biology and neuroscience. In such model systems, there is a recognition that it is not fully possible to study all possible variations, but that the knowledge gained from systematic study of a few carefully chosen representative systems can provide important insights for the field.

### Characteristics of model task paradigms

What would be the characteristics of model task paradigms used for motor learning? Although a number of factors may be involved in determining this (e.g., whether the timescale of learning is feasible, whether the difficulty is appropriate so that a majority of the participants can learn the task etc.), our view is that a task chosen for a model paradigm would have to score high at least on two dimensions: (a) relevance and (b) replication.

The ‘Relevance’ dimension measures how well the paradigm addresses the scientific question of interest. Tasks that score high on relevance would include consideration of features such as face validity (e.g., how well does the task represent the motor learning process of interest), high internal validity (the ability to tightly control extraneous factors in the experiment) and high external validity (the ability to generalize to other contexts) including ecological validity (the ability to generalize to learning of real-world tasks). It is important to note that many of these factors may be unknown at the time of the task development—so there is a need for domain expertise and qualitative judgment in determining relevance. Given that motor learning likely involves distinct types of processes in the different categories specified in Table 1, it is likely that model tasks for each of these processes would also have to be distinct to directly address these specific processes (i.e. a model task for adaptation would be different from a model task for sequence learning or one for variability). Importantly though, by specifying these model tasks at the level of these processes associated with learning (which are relatively few in number), researchers would be capable of reusing the tasks to address multiple research questions. For example, a model task of sequence learning could be used to address multiple varied research questions such as the role of sleep and consolidation (i.e. whether sleep enhances learning), contextual interference (i.e. whether random practice is superior to blocked practice), or the effect of self-controlled practice (i.e. whether having control of the practice sequence is superior to randomly practicing sequences).

The ‘Replication’ dimension measures how easy it is to replicate the paradigm in different labs, with access to potentially different resources. Tasks that score high on replication

would involve tasks with low reliance on specialized equipment while still allowing high measurement precision. For example, for a task involving control of variability, an underhand throw to a target would score higher on the replication dimension compared to a golf putt because it does not require a specific putter nor is it affected by environmental factors such as the friction between the ball and the surface. Similarly, a model task paradigm using inexpensive tools like webcam-based marker tracking (Krishnan et al. 2015) or markerless video tracking (Mathis et al. 2018) would score higher on the replication dimension than tasks requiring the use of specialized expensive equipment because it is likely to be more easily replicated in more labs with access to fewer resources.

Once an appropriate task is identified, the next step in making it a ‘model task paradigm’ is to ensure sufficient transparency that other researchers can replicate and build off these results. This involves two major steps—(a) the tasks are specified in enough detail that other groups can replicate the tasks as closely as possible, and (b) the data and analyses from these experiments are shared in a public repository so that the results from prior experiments can be combined and compared with results from future experiments. Practical guidelines for sharing of methods and data have been extensively reviewed in other domains (Gorgolewski and Poldrack 2016; Klein et al. 2018). Specifically in terms of motor learning, because of the richness and complexity of behavior possible, a particularly relevant solution is the use of video to help other researchers replicate the procedure more closely (Gilmore and Adolph 2017).

### Advantages of using model task paradigms

The use of model task paradigms directly addresses the challenges raised in the previous section. First, from a theoretical standpoint, model task paradigms permit direct replications which increases the likelihood of finding effects that are robust. Second, by adopting a ‘replicate and extend’ strategy (i.e. the experiment involves direct replication of a previous experiment but also collects data on some new parameter values), data from the first few studies would effectively yield ‘dose–response curves’ (such as those shown in Fig. 2d) that can provide important information about designing task parameter values for experimental manipulations. In fact, the use of model task paradigms opens the door for large scale studies across the globe that multiple labs can collaborate on—see for e.g. Psych Science Accelerator (Moshontz et al. 2018). These approaches may allow answering questions with large sample sizes that are currently not being investigated (e.g. individual differences) because they are beyond the scope of a single lab. Third, the presence of openly available data on a single task paradigm can produce more reliable estimates of effect

sizes, and also facilitate discussion on what theoretically meaningful effect sizes are. This information, analogous to the minimally clinically important difference (MCID) used in rehabilitation studies, is critical to making the distinction between ‘statistically significant’ and ‘meaningful’ results. Finally, using model task paradigms across different types of manipulations will allow direct comparisons in terms of raw effect sizes between different types of practice strategies, allowing practitioners with a good understanding of the relative utility of these strategies.

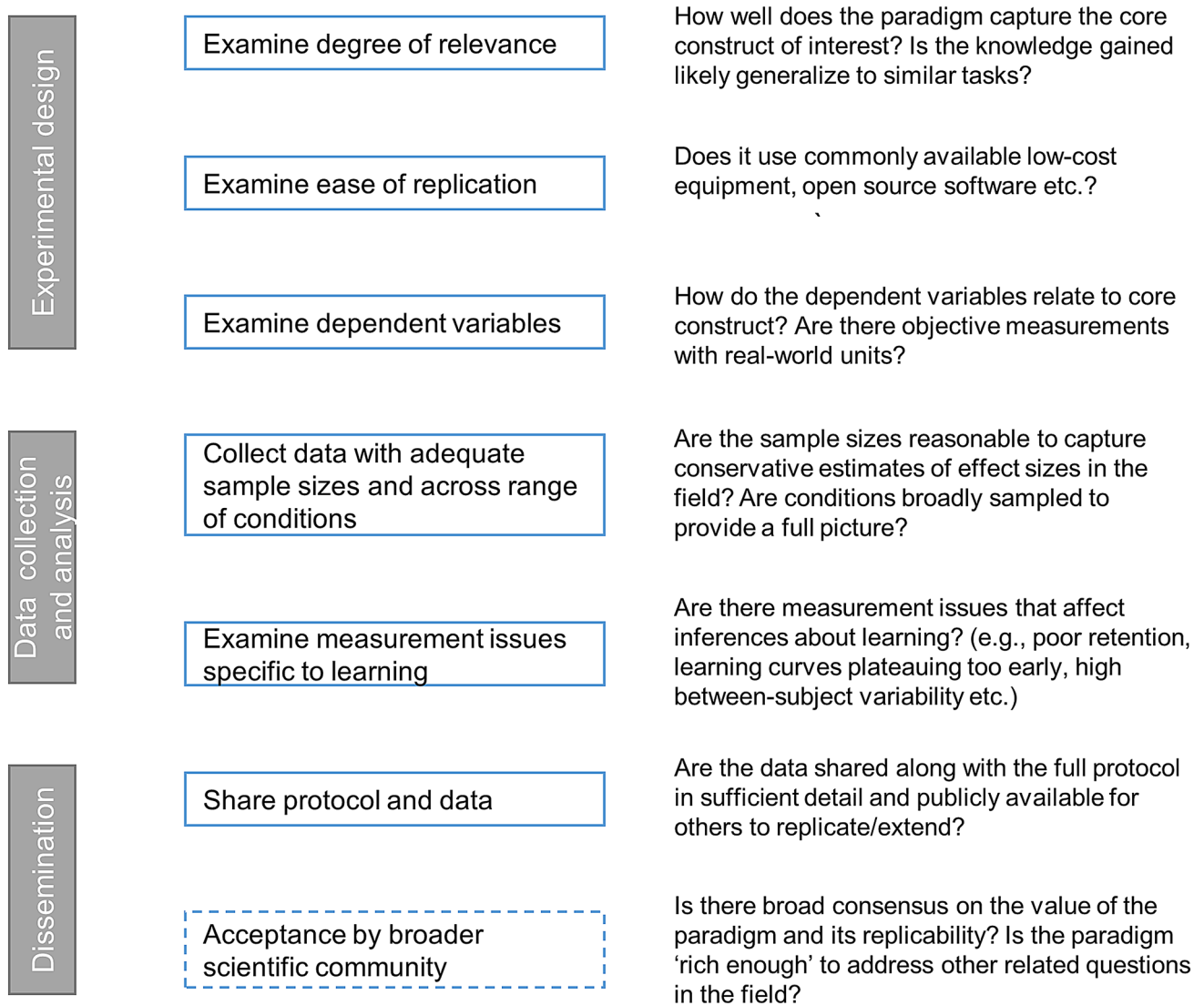
Beyond addressing these challenges, another feature of using model task paradigms is that they can effectively constrain ‘researcher degrees of freedom’ (Simmons et al. 2011). Although the term has been used to describe how undisclosed flexibility in data collection and analyses (such as flexibility in sample size or choosing among dependent variables) can make anything look ‘statistically significant’, the same issue also arises in the context of flexibility in task paradigms. For example, early studies on the role of augmented information in motor learning often chose task paradigms with extremely poor intrinsic information, such as drawing a line of a specified length. As a result, the role of augmented information was overrated in these tasks because it was often the only way participants could know what the task goal was (Swinnen 1996). Relatedly, the measurement of learning in these contexts has also used somewhat artificial situations such as No-KR tests, which often involve blindfolding participants from seeing the natural outcome of their actions (Russell and Newell 2007). Such criticisms have raised an important question of how relevant principles derived from simple tasks are to real-world learning (Wulf and Shea 2002; Russell and Newell 2007). Because model task paradigms are common only to broad themes in motor learning, and not at the level of individual research questions, they can effectively constrain flexibility in ‘tweaking’ of the task (intentional or unintentional) because the task and analyses are largely fixed in advance.

Last but not least, the use of model task paradigms also allows ‘data-driven’ discovery that could complement the dominant ‘hypothesis-driven’ approach in motor learning. The availability of relatively large data sets on a few standardized tasks could yield answers to questions that were not originally the focus of the work. An example of this in the motor learning has been the DREAM database (Walker and Kording 2013). Originally established as a collection of different experiments on reaching, data from these experiments were subsequently used to address a question about variability and rate of motor learning (He et al. 2016). In addition, these large data sets can also help serve to generate and test new theories or models of learning, as any new proposed theory or model at least needs to adequately accommodate for these data before making other testable predictions for future experiments.

The key steps involved in developing a model task are illustrated in Fig. 3. To demonstrate this with an example, consider the underarm ball toss to a target as a model task for learning to control variability (Rossum and Bootsma 1989). First, this task scores relatively high on relevance (it has good face validity because throwing the ball accurately to the target requires control of motor variability, internal validity can be increased through control of extraneous factors, and it also likely has ecological validity given several real-world motor learning tasks like the basketball free throw or golf putting require control of variability). Second, this task also scores relatively high on replication (because the only implement being used is a ball, such a paradigm is easy to replicate in any lab without the need for expensive or specialized equipment). Third, to make data in this task useful for other researchers, the dependent variable of task performance would have to be measured in ‘real-world units’—for e.g., the error from the target would need to be measured in centimeters (instead of a points scale for example). Fourth, initial studies using the task would aim to establish learning under a range of conditions involving variations of experimental parameters—for example, varying target distances or the amount of practice. Sample sizes for these initial studies may rely on broad ‘thumb rules’ for effect sizes (such as Cohen’s  $d$  of 0.5 etc.). Fifth, the data then have to be examined for how well they can be used to make inferences about the underlying learning—e.g., does task performance plateau too quickly, is the learning too variable between subjects, is the learning retained over a period of time? These questions will depend on what the underlying question of interest is—for example, high between-subject variability may actually be desirable if the goal is to examine individual differences. Sixth, the protocol and data is then deposited in a public repository (e.g., Open Science Framework) that is available for other researchers to use. The last and final step is how other scientists in the field perceive the proposed task—if the community is convinced of the utility of the proposed task to examine the motor learning process of interest, this task is adopted for further experiments and becomes a ‘model task’.

Once a model task is established, subsequent studies could then leverage this information in different ways. For example, one could start getting into manipulating practice strategies through dose–response studies. Using the underhand throw example, a study on variable practice could use multiple groups with wide range of practice variations (instead of the conventional two group design) to examine in what parameter range the ‘strongest’ effect of variable practice occurs in this task. Moreover, the learning curves established in the initial studies would also be informative in making magnitudes of effects more interpretable in terms of the time scale of learning (Day et al. 2018)—for e.g., if two groups differ by a throwing





**Fig. 3** A roadmap for constructing model tasks. At each stage of research (experimental design, data collection and analysis, and dissemination), the key steps and examples of associated questions that may need to be addressed are highlighted. The final step “acceptance by broader scientific community” is not undertaken directly by the experimenter but indicates how a proposed task eventually becomes

a model task. It is important to note that these processes and questions are not intended to be an exhaustive list for every context but rather provide a guide for such decisions. Ultimately, the goal of these model tasks is to enable work to be combined across multiple studies and labs in a way that can establish robust findings

error of 2 cm at the end of practice, that could be potentially translated into something like ‘using this practice strategy produced an improvement that would normally take 100 additional trials of practice’. Finally, because data are collected under the same standard conditions and shared across studies, it would become easier for future studies to determine effect sizes more precisely, which would then lead to more efficient sample sizes and a robust base of evidence for findings.

### Costs to using model task paradigms

From a practical standpoint, there is a cost in terms of the initial effort involved in developing model task paradigms compared to the status quo. Potential factors that drive this effort are (a) more careful consideration of tasks, (b) precise specification of associated measurement, (c) the use of larger sample sizes with more groups, and (d) the effort involved in making data and resources openly available. However,

we think that benefits are cumulative, with subsequent studies becoming much easier to plan and execute because it would allow investigators to skip over repeated pilot-testing phases, and use previously published data to make informed estimates about new experiments. Second, there is a potential risk of duplication if two research groups run the same experiment using the same model task. However, with the rising popularity of pre-prints and formats such as registered reports which allow for ‘results-blind’ acceptance (Caldwell et al. 2019; Chambers 2013), we believe that such concerns can be overcome.

More broadly for the field, there is a potential concern that model task paradigms may narrow the impact or generality of the field by decreasing the diversity of phenomena being studied. This concern has been expressed in the context of model organisms in biology (Yartsev 2017) and it is important not to fall back on the ‘easy’ route of only studying questions that can be answered using existing model task paradigms. At the level of the individual researcher, an overreliance on model tasks may hamper creativity and limit new discoveries. However, given that motor learning is currently at the other extreme with excessive fragmentation, we think that this concern, is at least for the moment, not a major one. In fact, as model task paradigms emerge for different themes, they may in fact actually help increase the diversity of problems studied by more clearly revealing which issues have received less attention, and provide opportunities for addressing such gaps through creative discovery. Moreover, model tasks themselves are not fixed but shaped by the scientific community—as some tasks reach a point of diminishing returns in terms of their utility, these could be replaced by other model tasks.

It is also perhaps worth re-iterating that model task paradigms are not meant to be a requirement for every experiment. Research questions at either extreme of the theoretical-applied spectrum are likely to continue to use customized tasks that suit their purpose. On the theoretical side, studies may involve a very specific manipulation (e.g., using a robotic exoskeleton to perturb a single joint) that requires the use of a task which does not fall into one of the model tasks. Similarly, on the applied side, there will always be a need for applied studies where the task itself is critical to the research question being answered (e.g., improving surgical technique). However, for the vast majority of studies in the middle of this spectrum, which have some flexibility in the choice of tasks, model task paradigms may provide a solution to the current level of fragmentation. These paradigms will also continue to evolve with greater theoretical understanding and improvements in measurement tools. Ultimately the success of any model task paradigm will depend on how other researchers in the field see its value, both in terms of the insights it generates, and in terms of how these insights generalize to the real world.

## Conclusion

In his highly influential paper on motor learning, Jack Adams (Adams 1971) criticized the use of real-world tasks saying that they resulted in ‘disconnected pockets of data’ that was unsuitable for the development of general scientific principles’. However, with the current level of fragmentation, we show that the same problem also exists even with laboratory tasks in motor learning. As a result, we believe that addressing this critical issue is vital for the field of motor learning. Many of the key steps outlined in Fig. 3 (standardizing protocols, dependent variables and open data sharing practices) have been recently discussed in other behavioral sciences for conducting large-scale multi-lab studies and clinical trials (Open Science Collaboration 2015; Adolph et al. 2012; Kwakkel et al. 2017; Frank et al. 2017), and it is our view that the field of motor learning may also benefit from such an effort.

Two related questions remain—(i) has this fragmentation always been the case in motor learning, and (ii) what are the underlying reasons for fragmentation? For the first question, we note that an early attempt to standardize tasks was undertaken in the “Learning Strategies Project” (Donchin 1989), which developed a computer game called Space Fortress (Mané and Donchin 1989) to allow direct comparisons between different learning strategies. Strikingly, the primary rationale stated for building a common computer game was task fragmentation as evidenced by the following quote “...it was quite evident that the diversity of paradigms and theoretical approaches within which the phenomena were studied, and the models tested, made it very difficult to compare results across studies” (Donchin 1989). Therefore, we do not think that the problem of task fragmentation is recent, although it is likely that the problem may have worsened in recent years as experimenters developed the tools to build their own hardware and software. For the second question of why such fragmentation exists, we can only speculate that there may be a number of factors that drive this fragmentation— incentives for novelty over replication, experimental ‘traditions’ that are handed down from mentors to graduate students, or even seemingly mundane issues like the limited availability of space or equipment, which increases the likelihood of creating new paradigms to shoehorn them to existing resources. Regardless of the underlying reasons, we suggest that it is time for research efforts to coalesce around a few model task paradigms for a more robust science that researchers in the field can build upon.

**Acknowledgements** The authors thank Dr. Chandramouli Krishnan for his comments on a previous version of this manuscript. The authors also thank Drs. Les Carlton and Mary Carlton for their comments, especially for pointing us to the work on the Learning strategies project.



This material is based upon work supported by the National Science Foundation under Grant No. 1823889 (RR) and a National Science Foundation Graduate Fellowship (ADT).

## References

- Abbas ZA, North JS (2018) Good-vs. poor-trial feedback in motor learning: the role of self-efficacy and intrinsic motivation across levels of task difficulty. *Learn Instr* 55:105–112
- Adams JA (1971) A closed-loop theory of motor learning. *J Mot Behav* 3:111–150
- Adolph KE, Gilmore RO, Freeman C, Sanderson P, Millman D (2012) Toward open behavioral science. *Psychol Inq* 23:244–247
- Aiken CA, Pan Z, Van Gemmert AWA (2017) The effects of a two-step transfer on a visuomotor adaptation task. *Exp Brain Res* 235:3459–3467
- Albers C, Lakens D (2018) When power analyses based on pilot data are biased: inaccurate effect size estimators and follow-up bias. *J Exp Soc Psychol* 74:187–195
- Baguley T (2009) Standardized or simple effect size: what should be reported? *Br J Psychol Lond Engl* 100(6):603–617
- Bahrick HP, Fitts PM, Briggs GE (1957) Learning curves; Facts or artifacts? *Psychol Bull* 54:255–268
- Bingham GP, Snapp-Childs W, Zhu Q (2018) Information about relative phase in bimanual coordination is modality specific (not amodal), but kinesthesia and vision can teach one another. *Hum Mov Sci* 60:98–106
- Bond KM, Taylor JA (2015) Flexible explicit but rigid implicit learning in a visuomotor adaptation task. *J Neurophysiol* 113:3836–3849
- Boyer ÉO, Bevilacqua F, Susini P, Hanneton S (2017) Investigating three types of continuous auditory feedback in visuo-manual tracking. *Exp Brain Res* 235:691–701
- Button KS, Ioannidis JPA, Mokrysz C, Nosek BA, Flint J, Robinson ESJ, Munafò MR (2013) Power failure: Why small sample size undermines the reliability of neuroscience. *Nat Rev Neurosci* 14:365–376
- Caldwell A, Vigotsky AD, Nuckols G, Boardley I, Schmidt J, Tenan M, Skarabot J, Radel R, Naughton M, Schoenfeld BJ, et al. (2019). Moving sport and exercise science forward: a call for the adoption of more transparent research practices (SportRxiv). <https://osf.io/fxe7a>. Accessed July 16 2019.
- Camerer CF, Dreber A, Holzmeister F, Ho T-H, Huber J, Johannesson M, Kirchler M, Nave G, Nosek BA, Pfeiffer T et al (2018) Evaluating the replicability of social science experiments in Nature and Science between 2010 and 2015. *Nat Hum Behav* 2:637
- Canaveral CA, Danion F, Berrigan F, Bernier P-M (2017) Variance in exposed perturbations impairs retention of visuomotor adaptation. *J Neurophysiol* 118:2745–2754
- Cardis M, Casadio M, Ranganathan R (2017) High variability impairs motor learning regardless of whether it affects task performance. *J Neurophysiol* 119:39–48
- Cardis M, Casadio M, Ranganathan R (2018) High variability impairs motor learning regardless of whether it affects task performance. *J Neurophysiol* 119:39–48
- Cattan E, Perrier P, Bérard F, Gerber S, Rochet-Capellan A (2018) Adaptation to visual feedback delays on touchscreens with hand vision. *Exp Brain Res* 236:3191–3201
- Chambers CD (2013) Registered reports: a new publishing initiative at cortex. *Cortex. J Devoted Study Nerv Syst Behav* 49:609–610
- Chambers CD (2017) *The seven deadly sins of psychology* (Princeton, NJ: Princeton University Press).
- Chauvel G, Maquestiaux F, Ruthruff E, Didierjean A, Hartley AA (2013) Novice motor performance: better not to verbalize. *Psychon Bull Rev* 20:177–183
- Chauvel G, Wulf G, Maquestiaux F (2015) Visual illusions can facilitate sport skill learning. *Psychon Bull Rev* 22:717–721
- Chien K-P, Chen S (2018) The influence of guided error-based learning on motor skills self-efficacy and achievement. *J Mot Behav* 50:275–284
- Chua L-K, Wulf G, Lewthwaite R (2018) Onward and upward: optimizing motor performance. *Hum Mov Sci* 60:107–114
- Couth S, Gowen E, Poliakoff E (2018) How does ageing affect grasp adaptation to a visual-haptic size conflict? *Exp Brain Res* 236:2173–2184
- Daou M, Lohse KR, Miller MW (2016) Expecting to teach enhances motor learning and information processing during practice. *Hum Mov Sci* 49:336–345
- Daou M, Hutchison Z, Bacelar M, Rhoads JA, Lohse KR, Miller MW (2019) Learning a skill with the expectation of teaching it impairs the skill's execution under psychological pressure. *J Exp Psychol Appl* 25:219–229
- Day KA, Leech KA, Roemmich RT, Bastian AJ (2018) Accelerating locomotor savings in learning: compressing four training days to one. *J Neurophysiol* 119:2100–2113
- de Barros JAC, Tani G, Corrêa UC (2017) Effects of practice schedule and task specificity on the adaptive process of motor learning. *Hum Mov Sci* 55:196–210
- Donchin E (1989) The learning strategies project: introductory remarks. *Acta Psychol (Amst)* 71:1–15
- Dotov D, Froese T (2018) Entraining chaotic dynamics: a novel movement sonification paradigm could promote generalization. *Hum Mov Sci* 61:27–41
- Doyon J, Korman M, Morin A, Dostie V, Tahar AH, Benali H, Karni A, Ungerleider LG, Carrier J (2009) Contribution of night and day sleep vs. simple passage of time to the consolidation of motor sequence and visuomotor adaptation learning. *Exp Brain Res* 195:15–26
- Dyer JF, Stapleton P, Rodger MWM (2017) Advantages of melodic over rhythmic movement sonification in bimanual motor skill learning. *Exp Brain Res* 235:3129–3140
- Fazeli D, Taheri H, Saberi Kakhki A (2017) Random versus blocked practice to enhance mental representation in golf putting. *Percept Mot Skills* 124:674–688
- Fialho JVAP, Tresilian JR (2017) Intercepting accelerated moving targets: effects of practice on movement performance. *Exp Brain Res* 235:1257–1268
- Frank C, Land WM, Schack T (2013) Mental representation and learning: the influence of practice on the development of mental representation structure in complex action. *Psychol Sport Exerc* 14:353–361
- Frank C, Land WM, Popp C, Schack T (2014) Mental representation and mental practice: experimental investigation on the functional links between motor memory and motor imagery. *PLoS ONE* 9:e95175
- Frank C, Land WM, Schack T (2016) Perceptual-cognitive changes during motor learning: the influence of mental and physical practice on mental representation, gaze behavior, and performance of a complex action. *Front Psychol* 6: <https://www.frontiersin.org/articles/10.3389/fpsyg.2015.01981/full>. Accessed Aug 27 2019.
- Frank MC, Bergelson E, Bergmann C, Cristia A, Floccia C, Gervain J, Hamlin JK, Hannon EE, Kline M, Levelt C et al (2017) A Collaborative approach to infant research: promoting reproducibility, best practices, and theory-building. *Infancy Off J Int Soc Infant Stud* 22:421–435
- Frank C, Kim T, Schack T (2018) Observational practice promotes action-related order formation in long-term memory: investigating action observation and the development of cognitive representation in complex motor action. *J Mot Learn Dev* 6:53–72

- French MA, Morton SM, Charalambous CC, Reisman DS (2018) A locomotor learning paradigm using distorted visual feedback elicits strategic learning. *J Neurophysiol* 120:1923–1931
- Gilmore RO, Adolph KE (2017) Video can make behavioural science more reproducible. *Nat Hum Behav* 1:0128
- Gorgolewski KJ, Poldrack RA (2016) A practical guide for improving transparency and reproducibility in neuroimaging research. *PLoS Biol* 14:e1002506
- Grand KF, Daou M, Lohse KR, Miller MW (2017) Investigating the mechanisms underlying the effects of an incidental choice on motor learning. *J Mot Learn Dev* 5:207–226
- Guadagnoli MA, Lee TD (2004) Challenge point: a framework for conceptualizing the effects of various practice conditions in motor learning. *J Mot Behav* 36:212–224
- Haarman JAM, Choi JT, Buurke JH, Rietman JS, Reenalda J (2017) Performance of a visuomotor walking task in an augmented reality training setting. *Hum Mov Sci* 56:11–19
- Hasson CJ, Zhang Z, Abe MO, Sternad D (2016) Neuromotor noise is malleable by amplifying perceived errors. *PLoS Comput Biol* 12:e1005044
- He K, Liang Y, Abdollahi F, Fisher Bittmann M, Kording K, Wei K (2016) The statistical determinants of the speed of motor learning. *PLoS Comput Biol* 12:e1005023
- Hebert E (2018) The effects of observing a learning model (or two) on motor skill acquisition. *J Mot Learn Dev* 6:4–17
- Hinkel-Lipsker JW, Hahn ME (2017) The effects of variable practice on locomotor adaptation to a novel asymmetric gait. *Exp Brain Res* 235:2829–2841
- Hinkel-Lipsker JW, Hahn ME (2018) Coordinative structuring of gait kinematics during adaptation to variable and asymmetric split-belt treadmill walking—a principal component analysis approach. *Hum Mov Sci* 59:178–192
- Holland P, Codol O, Galea JM (2018) Contribution of explicit processes to reinforcement-based motor learning. *J Neurophysiol* 119:2241–2255
- Itaguchi Y, Fukuzawa K (2018) Influence of speed and accuracy constraints on motor learning for a trajectory-based movement. *J Mot Behav* 50:653–663
- Jalali R, Miall RC, Galea JM (2017) No consistent effect of cerebellar transcranial direct current stimulation on visuomotor adaptation. *J Neurophysiol* 118:655–665
- Jiang W, Yuan X, Yin C, Wei K (2018) Visuomotor learning is dependent on direction-specific error saliency. *J Neurophysiol* 120:162–170
- Kaipa R, Kaipa RM (2018) Role of constant, random and blocked practice in an electromyography-based oral motor learning task. *J Mot Behav* 50:599–613
- Kaipa R, Robb M, Jones R (2017) The effectiveness of constant, variable, random, and blocked practice in speech-motor learning. *J Mot Learn Dev* 5:103–125
- Karlinsky A, Hodges NJ (2018) Dyad practice impacts self-directed practice behaviors and motor learning outcomes in a contextual interference paradigm. *J Mot Behav* 50:579–589
- Kearney PE (2015) A distal focus of attention leads to superior performance on a golf putting task. *Int J Sport Exerc Psychol* 13:371–381
- Kerr R, Booth B (1978) Specific and varied practice of motor skill. *Percept Mot Skills* 46:395–401
- Kim T, Frank C, Schack T (2017) A Systematic investigation of the effect of action observation training and motor imagery training on the development of mental representation structure and skill performance. *Front Hum Neurosci*. <https://link-gategroup-com.proxy2.cl.msu.edu/apps/doc/A511188490/AONE?sid=lms>. Accessed Aug 27 2019.
- Kim JH, Han JK, Han DH (2018) Training effects of interactive metronome® on golf performance and brain activity in professional woman golf players. *Hum Mov Sci* 61:63–71
- Kimura T, Kaneko F, Nagahata K, Shibata E, Aoki N (2017) Working memory training improves dual-task performance on motor tasks. *J Mot Behav* 49:388–397
- Klein O, Hardwicke TE, Aust F, Breuer J, Danielsson H, Mohr AH, Ijzerman H, Nilsson G, Vanpaemel W, Frank MC (2018) A Practical guide for transparency in psychological science. *Collabra Psychol* 4:20
- Krajenbrink H, van Abswoude F, Vermeulen S, van Cappellen S, Steenbergen B (2018) Motor learning and movement automatization in typically developing children: the role of instructions with an external or internal focus of attention. *Hum Mov Sci* 60:183–190
- Krakauer JW, Mazzoni P (2011) Human sensorimotor learning: adaptation, skill, and beyond. *Curr Opin Neurobiol* 21:636–644
- Krause D, Brüne A, Fritz S, Kramer P, Meisterjahn P, Schneider M, Sperber A (2014) Learning of a golf putting task with varying contextual interference levels induced by feedback schedule in novices and experts. *Percept Mot Skills* 118:384–399
- Krause D, Agethen M, Zobe C (2018) Error feedback frequency affects automaticity but not accuracy and consistency after extensive motor skill practice. *J Mot Behav* 50:144–154
- Krishnan C, Washabaugh EP, Seetharaman Y (2015) A low cost real-time motion tracking approach using webcam technology. *J Biomech* 48:544–548
- Kumar N, Kumar A, Sonane B, Mutha PK (2018) Interference between competing motor memories developed through learning with different limbs. *J Neurophysiol* 120:1061–1073
- Kwakkel G, Lannin NA, Borschmann K, English C, Ali M, Churilov L, Saposnik G, Winstein C, van Wegen EE, Wolf SL et al (2017) Standardized measurement of sensorimotor recovery in stroke trials: consensus-based core recommendations from the stroke recovery and rehabilitation roundtable. *Int J Stroke Off J Int Stroke Soc* 12:451–461
- Lakens D, Scheel AM, Isager PM (2018) Equivalence testing for psychological research: a tutorial. *Adv Methods Pract Psychol Sci* 1:259–269
- Land WM, Frank C, Schack T (2014) The influence of attentional focus on the development of skill representation in a complex action. *Psychol Sport Exerc* 15:30–38
- Lawrence GP, Cassell VE, Beattie S, Woodman T, Khan MA, Hardy L, Gottwald VM (2014) Practice with anxiety improves performance, but only when anxious: evidence for the specificity of practice hypothesis. *Psychol Res* 78:634–650
- Lee TD, Genovese ED (1989) Distribution of practice in motor skill acquisition: different effects for discrete and continuous tasks. *Res Q Exerc Sport* 60:59–65
- Leow L-A, Gunn R, Marinovic W, Carroll TJ (2017) Estimating the implicit component of visuomotor rotation learning by constraining movement preparation time. *J Neurophysiol* 118:666–676
- Levac D, Driscoll K, Galvez J, Mercado K, O'Neil L (2017) OPTIMAL practice conditions enhance the benefits of gradually increasing error opportunities on retention of a stepping sequence task. *Hum Mov Sci* 56:129–138
- Lewthwaite R, Chiviawsky S, Drews R, Wulf G (2015) Choose to move: the motivational impact of autonomy support on motor learning. *Psychon Bull Rev* 22:1383–1388
- Lin T-H, Denomme A, Ranganathan R (2018) Learning alternative movement coordination patterns using reinforcement feedback. *Exp Brain Res*. <https://doi.org/10.1007/s00221-018-5227-1>
- Lo AC, Guarino PD, Richards LG, Haselkorn JK, Wittenberg GF, Federman DG, Ringer RJ, Wagner TH, Krebs HI, Volpe BT et al (2010) Robot-assisted therapy for long-term upper-limb impairment after stroke. *N Engl J Med* 362:1772–1783

- Lohse K, Buchanan T, Miller M (2016) Underpowered and overworked: problems with data analysis in motor learning studies. *J Mot Learn Dev* 4:37–58
- LoJacono CT, MacPherson RP, Kuznetsov NA, Raisbeck LD, Ross SE, Rhea CK (2018) Obstacle crossing in a virtual environment transfers to a real environment. *J Mot Learn Dev* 6:234–249
- Maeda RS, McGee SE, Marigold DS (2016) Consolidation of visuomotor adaptation memory with consistent and noisy environments. *J Neurophysiol* 117:316–326
- Mané A, Donchin E (1989) The space fortress game. *Acta Psychol (Amst)* 71:17–22
- Mansfield A, Aqai A, Fraser JE, Rajachandrakumar R, Lakhani B, Patterson KK (2017) Can augmented feedback facilitate learning a reactive balance task among older adults? *Exp Brain Res* 235:293–304
- Marchal-Crespo L, Rappo N, Riener R (2017) The effectiveness of robotic training depends on motor task characteristics. *Exp Brain Res* 235:3799–3816
- Mathis A, Mamidanna P, Cury KM, Abe T, Murthy VN, Mathis MW, Bethge M (2018) DeepLabCut: markerless pose estimation of user-defined body parts with deep learning. *Nat Neurosci* 21:1281–1289
- McGregor HR, Gribble PL (2017) Functional connectivity between somatosensory and motor brain areas predicts individual differences in motor learning by observing. *J Neurophysiol* 118:1235–1243
- McGregor HR, Cashaback JGA, Gribble PL (2018) Somatosensory perceptual training enhances motor learning by observing. *J Neurophysiol* 120:3017–3025
- McKenna E, Bray LCJ, Zhou W, Joiner WM (2017) The absence or temporal offset of visual feedback does not influence adaptation to novel movement dynamics. *J Neurophysiol* 118:2483–2498
- Meira CM, Fairbrother JT (2018) Ego-oriented learners show advantage in retention and transfer of balancing skill. *J Mot Learn Dev* 6:209–219
- Melendez-Calderon A, Tan M, Bittmann MF, Burdet E, Patton JL (2017) Transfer of dynamic motor skills acquired during isometric training to free motion. *J Neurophysiol* 118:219–233
- Milner TE, Firouzimehr Z, Babadi S, Ostry DJ (2018) Different adaptation rates to abrupt and gradual changes in environmental dynamics. *Exp Brain Res* 236:2923–2933
- Mischel, W. (2008). The Toothbrush Problem. *APS Obs.* 21. Available at: <https://www.psychologicalscience.org/observer/the-toothbrush-problem> [Accessed April 4, 2019].
- Moshontz H, Campbell L, Ebersole CR, IJerman H, Urry HL, Forscher PS, Grahe JE, McCarthy RJ, Musser ED, Antfolk J et al (2018) The Psychological Science Accelerator: advancing psychology through a distributed collaborative network. *Adv Methods Pract. Psychol. Sci.* 1:501–515
- Munzert J, Maurer H, Reiser M (2014) Verbal-motor attention-focusing instructions influence kinematics and performance on a golf-putting task. *J Mot Behav* 46:309–318
- Navarro M, van der Kamp J, Schor P, Savelsbergh GJP (2018) Implicit learning increases shot accuracy of football players when making strategic decisions during penalty kicking. *Hum Mov Sci* 61:72–80
- Neville K-M, Trempe M (2017) Serial practice impairs motor skill consolidation. *Exp Brain Res* 235:2601–2613
- Newell, K.M. (1986). Constraints on the development of coordination. In: *Motor development in children: Aspects of coordination and control*, M. G. Wade and H. T. A. Whiting, eds. (Nijhoff).
- Newell KM (1989) On task and theory specificity. *J Mot Behav* 21:92–96
- Newell KM (1991) Motor skill acquisition. *Annu Rev Psychol* 42:213–237
- Nunes ME, de S, Correa UC, Souza MGTX, de Basso L, Coelho DB, Santos S (2019) No improvement on the learning of golf putting by older persons with self-controlled knowledge of performance. *J Aging Phys Act* 27:300–308
- Ong NT, Hodges NJ (2018) Balancing our perceptions of the efficacy of success-based feedback manipulations on motor learning. *J Mot Behav* 50:614–630
- Open Science Collaboration (2015) Estimating the reproducibility of psychological science. *Science* 349:aac716
- Overduin SA, Richardson AG, Lane CE, Bizzi E, Press DZ (2006) Intermittent practice facilitates stable motor memories. *J Neurosci* 26:11888–11892
- Pacheco MM, Newell KM (2018) Learning a specific, individual and generalizable coordination function: evaluating the variability of practice hypothesis in motor learning. *Exp Brain Res* 236:3307–3318
- Palmer K, Chiviawsky S, Wulf G (2016) Enhanced expectancies facilitate golf putting. *Psychol Sport Exerc* 22:229–232
- Panzer S, Kennedy D, Wang C, Shea CH (2018) The simplest acquisition protocol is sometimes the best protocol: performing and learning a 1:2 bimanual coordination task. *Exp Brain Res* 236:539–550
- Pashler H, Harris CR (2012) Is the replicability crisis overblown? Three arguments examined. *Perspect Psychol Sci* 7:531–536
- Pinder RA, Davids K, Renshaw I, Araújo D (2011) Representative learning design and functionality of research and practice in sport. *J Sport Exerc Psychol* 33:146–155
- Raisbeck LD, Diekfuss JA (2017) Verbal cues and attentional focus: a simulated target-shooting experiment. *J Mot Learn Dev* 5:148–159
- Reilly KJ, Pettibone C (2017) Vowel generalization and its relation to adaptation during perturbations of auditory feedback. *J Neurophysiol* 118:2925–2934
- Ring C, Cooke A, Kavussanu M, McIntyre D, Masters R (2015) Investigating the efficacy of neurofeedback training for expediting expertise and excellence in sport. *Psychol Sport Exerc* 16:118–127
- Rossum JHAV, Bootsma RJ (1989) The underarm throw for accuracy in children. *J Sports Sci* 7:101–112
- Ruffino C, Papaxanthis C, Lebon F (2017) The influence of imagery capacity in motor performance improvement. *Exp Brain Res* 235:3049–3057
- Russell DM, Newell KM (2007) On No-KR tests in motor learning, retention and transfer. *Hum Mov Sci* 26:155–173
- Schmitz G, Bock OL (2017) Properties of intermodal transfer after dual visuo- and auditory-motor adaptation. *Hum Mov Sci* 55:108–120
- Seidler RD, Gluskin BS, Greeley B (2016) Right prefrontal cortex transcranial direct current stimulation enhances multi-day savings in sensorimotor adaptation. *J Neurophysiol* 117:429–435
- Shuggi IM, Shewokis PA, Herrmann JW, Gentili RJ (2018) Changes in motor performance and mental workload during practice of reaching movements: a team dynamics perspective. *Exp Brain Res* 236:433–451
- Simmons JP, Nelson LD, Simonsohn U (2011) False-positive psychology: undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychol Sci* 22:1359–1366
- Sobierajewicz J, Przekoracka-Krawczyk A, Jaśkowski W, van der Lubbe RHJ (2017) How effector-specific is the effect of sequence learning by motor execution and motor imagery? *Exp Brain Res* 235:3757–3769
- Song Y, Smiley-Oyen AL (2017) Probability differently modulating the effects of reward and punishment on visuomotor adaptation. *Exp Brain Res* 235:3605–3618
- Swinnen SP (1996) Information feedback for motor skill learning: a review. In: Zelaznik HN (ed) *Advances in motor learning and control*. Human Kinetics, Champaign, IL, pp 37–66

- Thorp EB, Kording KP, Mussa-Ivaldi FA (2016) Using noise to shape motor learning. *J Neurophysiol* 117:728–737
- van Ginneken WF, Poolton JM, Capio CM, van der Kamp J, Choi CSY, Masters RSW (2018) Conscious control is associated with freezing of mechanical degrees of freedom during motor learning. *J Mot Behav* 50:436–456
- Vine S, Moore L, Cooke A, Ring C, Wilson M (2013) Quiet eye training: a means to implicit motor learning quiet eye training: a means to implicit motor learning. *Int J Sport Psychol* 44:367
- Walker B, Kording K (2013) The database for reaching experiments and models. *PLoS ONE* 8:e78747
- Wulf G, Shea CH (2002) Principles derived from the study of simple skills do not generalize to complex skill learning. *Psychon Bull Rev* 9:185–211
- Wulf G, Iwatsuki T, Machin B, Kellogg J, Copeland C, Lewthwaite R (2018) Lassoing skill through learner choice. *J Mot Behav* 50:285–292
- Yartsev MM (2017) The emperor's new wardrobe: rebalancing diversity of animal models in neuroscience research. *Science* 358:466–469
- Yen S-C, Olsavsky LC, Cloonan CM, Llanos AR, Dwyer KJ, Nabian M, Farjadian AB (2018) An examination of lower limb asymmetry in ankle isometric force control. *Hum Mov Sci* 57:40–49
- Yokoi A, Bai W, Diedrichsen J (2016) Restricted transfer of learning between unimanual and bimanual finger sequences. *J Neurophysiol* 117:1043–1051
- Zhu FF, Yeung AY, Poolton JM, Lee TMC, Leung GKK, Masters RSW (2015) Cathodal transcranial direct current stimulation over left dorsolateral prefrontal cortex area promotes implicit motor learning in a golf putting task. *Brain Stimulat* 8:784–786
- Ziv G, Lidor R (2015) Focusing attention instructions, accuracy, and quiet eye in a self-paced task—an exploratory study. *Int J Sport Exerc Psychol* 13:104–120
- Zwaan RA, Etz A, Lucas RE, Donnellan MB (2018) Making replication mainstream. *Behav Brain Sci* 41:e120

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.