Routledge
Taylor & Francis Group

Human Performance

ISSN: (Print) (Online) Journal homepage: www.tandfonline.com/journals/hhup20

A Comparison of Individual and Team Skill Acquisition, Retention (Decay), and Reacquisition Using a Synthetic Task Environment

Winfred Arthur Jr, Andrew M. Naber, Gonzalo J. Muñoz, Nathanael L. Keiser, Olabisi A. Atoba, Inchul Cho & Winston Bennett Jr

To cite this article: Winfred Arthur Jr, Andrew M. Naber, Gonzalo J. Muñoz, Nathanael L. Keiser, Olabisi A. Atoba, Inchul Cho & Winston Bennett Jr (2025) A Comparison of Individual and Team Skill Acquisition, Retention (Decay), and Reacquisition Using a Synthetic Task Environment, Human Performance, 38:1, 43-60, DOI: 10.1080/08959285.2024.2436171

To link to this article: https://doi.org/10.1080/08959285.2024.2436171

HUMAN PERFORMANCE
2025, VOL. 38, NO. 1, 43–60
https://doi.org/10.1080/08959285.2024.2436171

A Comparison of Individual and Team Skill Acquisition, Retention (Decay), and Reacquisition Using a Synthetic Task Environment

Winfred Arthur Jr[a], Andrew M. Naber[a,b], Gonzalo J. Muñoz[c], Nathanael L. Keiser[a,b], Olabisi A. Atoba[a], Inchul Cho[a,d], and Winston Bennett Jr[e]

[a]Texas A&M University; [b]US Army Research Institute for the Behavioral and Social Sciences (ARI); [c]School of Management, Pontificia Universidad Católica de Chile; [d]University of North Georgia; [e]U.S. Air Force

ABSTRACT
Skill decay is the decrement in performance on acquired knowledge and skills after a period of nonuse. Although there is a fair amount of work on individual skill decay, the literature on team-level skill decay is very limited. Thus, using a synthetic task environment, the objective of this study was to compare individual and team performance in terms of skill acquisition, retention (decay), and reacquisition. Eighty-one individuals in 27 three-person teams were trained to perform a complex computer-based simulation. The initial acquisition phase comprised a 2-day, 5-hour training protocol. After an average 73.33-day nonuse interval ($SD = 30.18$), participants returned for a 2-hour reacquisition session. Participants completed the performance task as both individuals and teams. The results indicated that, compared to individual performance, team performance improved faster during the skill *acquisition* phase. However, unlike individual performance, the amount of *decay* for team performance was dependent on whether the first retention performance session for team members was performed as individuals—where trainees had the opportunity to perform all components of the task—or as a team, with the latter resulting in the most decay. By extension, these results highlight the impact of a relatively brief individual post-retention (nonuse) interval training to mitigate team skill decay. In short, our results indicate that a little individual practice can go a long way toward maintaining team retention performance, but in contrast, a little team practice does not do the same for individual retention performance.

Skill decay refers to observed decrements in acquired skills (or knowledge) after a period of nonuse, that is, a retention interval (Arthur et al., 1998). Skill decay is particularly salient in situations where individuals and teams do not receive refresher training, cannot regularly perform acquired skills, or only perform these skills after extended nonuse/retention intervals where they are expected to perform at full proficiency. While there has been a continued interest in skill and knowledge decay in the both the scholarly and applied literatures, most of this has focused on individuals with very limited attention to teams. Consistent with this, a detailed search of the pertinent literature from 2012–2024 identified 67 individual skill decay papers, but only three team skill decay papers (see Figure 1[1]), all of which were chapters in edited volumes; lending support to the summary statement that there has been a dearth of discussion about or research on team skill decay or loss.

CONTACT Winfred Arthur, Jr. w-arthur@tamu.edu Department of Psychological and Brain Sciences, Texas A&M University, College Station, TX 4235 TAMU

This article has been republished with minor changes. These changes does not impact on the academic content of the article.
[1]We conducted a literature search of relevant databases (i.e., PsycINFO, Academic Source Complete, and Business Source Complete) to identify published and unpublished papers encompassing January 2012 – June 2024 (inclusive) using "skill decay" as the search term. A complete reference list of the papers identified in this search is included in the online supplementary materials.
Supplemental data for this article can be accessed online at https://doi.org/10.1080/08959285.2024.2436171.

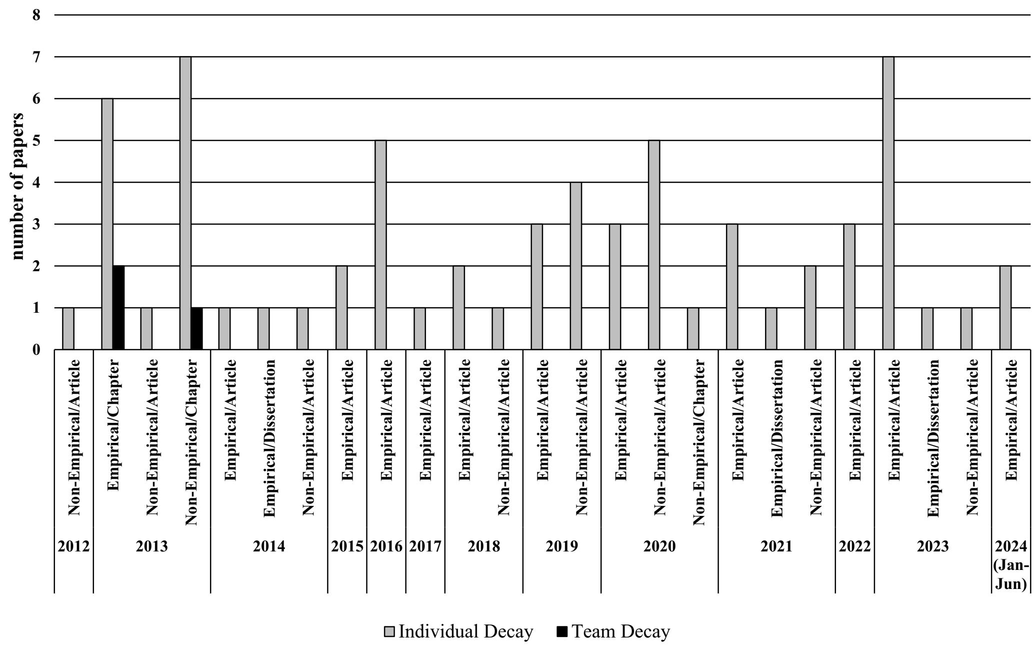

Figure 1. Frequency plot of published and unpublished papers on individual and team skill decay in the past 12 years (2012–2024).

The limited attention to team skill decay is in sharp contrast to the rich and voluminous literature on team training and team skill acquisition (e.g., Hughes et al., 2016; Jarrett et al., 2016; Keiser & Arthur, 2021, 2022; Salas et al., 2008b; Villado & Arthur, 2013). Consequently, an examination of team skill decay and the equally important issue of reacquisition in an effort to begin to fill these gaps in the literature is warranted for a host of reasons. First, as noted by Schmidt and Bjork (1992), acquisition, retention/loss, and reacquisition are separate phenomena that may yield different interpretations about the effectiveness of training interventions. That is, consonant with the concept of desirable difficulties (Bjork, 1994; Bjork & Bjork, 2011), training interventions that are effective in terms of acquisition may not necessarily be the most effective in terms of long-term retention and/or reacquisition and vice versa. Thus, the team skill *acquisition* literature may not necessarily generalize to or permit inferences about team skill *decay and reacquisition*.

Second, the skill decay literature is predominantly characterized by the use of simple tasks such as word lists, simple facts, and vocabulary words as examples (Arthur et al., 1998; Cepeda et al., 2008; Sobel et al., 2011). Consequently, given the noted differences between simple and complex tasks, the assumption that the decay or retention of cognitively complex decision-making tasks is similar to that of simple tasks may be unfounded. For instance, complex tasks may be more resistant to skill decay than simple tasks due to the deeper processing and learning engendered during training (Wang et al., 2013). Similar observations about the importance of distinguishing simple from complex tasks have been made by Klostermann et al. (2022) in the context of skill decay in high-risk industries. Furthermore, as noted by Arthur and Day (2013a), compared to simple tasks, complex tasks involve more task parameters such that complexity may covary with longer performance episodes. So, to fully capture performance on complex tasks requires longer performance episodes to allow all task parameters to run their course. Longer retention tests in turn provide individuals and teams with more opportunity to reacquaint themselves with the task. Consequently, complex tasks may appear more resistant to skill loss because they provide inherent opportunities to reacquire skill on the retention test.

In spite of this, since Arthur et al.'s (1998) meta-analysis (see Wang et al., 2013 as well), there have been a relatively small number of studies that have examined the decay of complex skills (e.g., Day et al., 2001; Frank & Kluge, 2018, 2019; Kluge & Frank, 2014; May & Kahnweiler, 2000; Muñoz et al., 2022; Sauer et al., 2000; Shebilske et al., 1999; Woollard et al., 2006), and the general pattern of results from these albeit relatively small number of studies indicate that the magnitude and rate of skill decay on complex tasks is quite different from that observed for simple tasks.

Third, although skill decay can also manifest in teams, the observed magnitude and patterns of skill decay in individuals may not necessarily generalize to teams. This is because the interdependence that characterizes teams—which "consist of two or more individuals who work interdependently, have specific role assignments, perform specific tasks, and interact and coordinate to achieve a common goal" (Arthur & Day, 2013b, p. 11)—may engender unique team-level processes that influence skill acquisition, decay/retention, and reacquisition in a manner that is not germane to individuals. For instance, teamwork processes may influence the way in which team members' knowledge is acquired, retrieved, and effectively implemented (Cooke et al., 2004). The presence of additional team-level cognitive mechanisms such as shared mental models (DeChurch & Mesmer-Magnus, 2010a, 2010b; Mohammed et al., 2010), transactive memory systems (Kozlowski & Ilgen, 2006), and cross-training (Marks et al., 2002) suggest that team-level processes might even compensate for any decay occurring at the individual level (Cooke et al., 2013). This perspective is consistent with Arthur et al.'s (2013) finding that teams perform better than individuals in terms of transfer to novel tasks. By the same token, the challenges of team processes may hinder teams' ability to perform effectively during initial skill acquisition, compared to the training of individuals. Thus, comparing individual and team skill decay and reacquisition is critical given its potential to advance the field of complex-skill training. However, with the exception of Arthur et al. (2013) and Cooke et al. (2013), we were unable to locate any published empirical or conceptual reports on *comparative* examinations of individual and team skill retention and reacquisition.

Arthur et al. (2013) describe and present a skill acquisition and retention study in which they compare individual- and team-level acquisition, retention, and transfer on a complex command-and-control simulation task with an 8-week nonuse/retention interval. Their results indicated that individuals and (three-person) teams displayed differing levels of skill acquisition and transfer, but similar levels of retention. In addition, both individuals and teams displayed smaller levels of skill decay after the 8-week nonuse/retention interval ($d = -0.11$ and -0.12, respectively) than researchers may be inclined to anticipate—a finding they attribute to the complex nature of the task. Finally, the spacing of practice (23-hour versus 2-minute intersession intervals), had a larger effect on individuals than teams. However, Arthur et al. used a between-subjects design in which the individual data collection was completed before the team data collection began. So, the use of a between-subject design coupled with the sequential collection of the individual and team data instead of simultaneously with trainees being randomly assigned to the team or individual conditions, pose reasonable queries about the equivalency of the team and individual conditions and the resultant data. The present study addresses these methodological concerns by using a within-subjects design.

In contrast to Arthur et al. (2013) who focused on taskwork, Cooke et al. (2013) focused on teamwork by investigating the retention of team coordination skills. Specifically, the results of a series of lab studies indicated that team coordination skills in a three-person unmanned aerial vehicle command-and-control task decayed after a nonuse period of 10 weeks or sooner, although the loss was short-lived. Furthermore, in examining the relative contributions of individual competency and team interaction skills to the prediction of team performance decay, Cooke et al. concluded that team performance decay is accounted for more by differences in team member interaction, than it is by individual competency. Hence, Cooke et al.'s results support the proposition that team performance is more than the sum of individual team member performance, and that team interaction processes play a major role in differences in team skill retention.

So, given the relative nascency of the team skill decay literature, the present study sought to compare individual and team skill acquisition, retention (decay), and reacquisition in the context of a synthetic task environment. Synthetic task environments (see Sitzmann, 2011) are tasks that simulate and model the cognitive, information processing, psychological, and social processes and demands that are present in operational environments. They are particularly well suited for training research because they can be configured to simulate a wide range of conditions and events, and also allow for the collection of detailed performance data. Examples of synthetic task environments in the literature include computer-based simulators and desktop trainers, such as Steel Beasts Pro PE (eSim Games, 2007; Jarrett et al., 2016), Unmanned Aerial Vehicle-Synthetic Task Environment (Cooke & Shope, 2005; Gorman & Cooke, 2011), Unreal Tournament (Epic Games, Inc, 2004; Hughes et al., 2013), and Space Fortress (Mane & Donchin, 1989; Schuelke et al., 2009). The present study used Crisis in the Kodiak: Oilrig Search and Rescue (Arthur et al., 2011), which is a prototypical example of the types of complex synthetic task environments used in lab-based team training research. In summary, consonant with the preceding review, the present study sought to answer the following question:

Research Question: Is team performance, in terms of (a) skill acquisition, (b) retention (decay), and (c) reacquisition in the context of a synthetic task environment, different or the same as that for individual performance?

Method

Participants

Participants were recruited from the research participant pool of the psychology department at a large southwest U.S. university. The study protocol was approved by the Institutional Review Board of the first author's university and adheres to the tenets of the Declaration of Helsinki. The sample comprised 81 individuals randomly assigned to 27 three-person teams[2] and was 71.60% female with a mean age of 18.79 years (*SD* = 1.15). A 3-point (1 = *novice* 2 = *average* and 3 = *expert*) single item was used to assess participants' video game experience. Only ten participants (12.35%) self-described as *expert* with the majority describing themselves as *average* (*n* = 39 48.15%) or *novice* (*n* = 32 39.51%) video game players. Participation in the acquisition phase of the study fulfilled a course requirement. However, to encourage participants to return to complete the remainder of the study after the scheduled 6-week nonuse interval, they were paid $30 for the reacquisition phase of the study. Furthermore, to motivate participants to perform well, they were also eligible to earn a monetary reward of $80, $40, or $20 (per team member) for teams that attained the three highest average team performance scores, respectively. Similarly, participants were also eligible to earn a monetary reward of $80, $40, or $20 (per individual) for individuals who attained the three highest average individual performance scores, respectively. Thus, participants could earn performance rewards for their individual and team performance, independently.

Measures

Performance task - Crisis in the Kodiak: Oilrig Search and Rescue (Crisis; Arthur et al., 2011)

Participants were trained to perform a dynamic, networked computer-based simulation. *Crisis* was developed to permit trainees to perform the task as either individuals or teams. It was also designed to not require any previous military or video game experience. Participants operated the simulator

[2]Whereas we had initially sought to run a larger number of teams, the final sample size was the result of the challenges encountered in recruiting three-person teams spanning three separate days of data collection (Figure 3). Consequently, to provide an important boundary condition to interpret the results, a post hoc power analysis was conducted. The results of this analysis indicated that the final sample (*n* = 27 teams) achieved a power of .82 to detect a statistically significant (*p* = .05) effect of *d* = 0.71 between team performance decay in two conditions. The effect size used for this post hoc power analysis is the most impactful finding observed in the present study; namely, the difference between team performance decay across two trial order conditions (individual mission followed by team mission, and vice versa).

through a command-and-control interface using networked desktop computers with two monitors, a keyboard, and a mouse. Participants performed the individual and team missions in the same room at their own computer stations. During team missions, participants communicated with each other via voice-activated microphones and headsets. A full and detailed description of this training task is reported in Arthur et al. (2011), and Figure 2 presents a screen capture of the game screen.

Crisis in the Kodiak Missions. Participants operated the simulator either collectively as a three-person team (taking on specialized roles), or as individuals (performing all roles simultaneously). *Crisis* is a disaster response simulation that was developed to include task, goal, and feedback interdependencies. Missions included the roles of oilrig workers, helicopter aviators, and boat captains tasked with responding to an off-shore oilrig explosion. When performing as a team, each team operated nine platforms (three for each role) to achieve two goals of shutting off oil valves and rescuing survivors, and each team member coordinated the three platforms that comprised his or her assigned role. Platforms had unique capabilities that were used individually and interactively to accomplish mission objectives. Each of these three types of platforms varied in its capacities on six different dimensions: (1) range of vision, (2) speed of movement, (3) type of movement (land, air, and sea), (4) type of capabilities (alternate methods for putting out fires and healing survivors), (5) range of capabilities, and (6) ability to pick up survivors. Each mission lasted 10 minutes. When performing as individuals, participants

Figure 2. Example screenshot of the *Crisis in the Kodiak: Oilrig Search and Rescue* simulation.

Note. This screenshot depicts *Crisis* from the perspective of a helicopter aviator during a team mission. A sequence of simultaneous engagements identified by red lines is occurring on the oilrig platform in this screenshot. On the right side, the helicopter aviator is using two helicopters (Helo 1 and Helo 3, respectfully) to extinguish structural fires on the oilrig. In the bottom left, the oilrig worker is using three workers (OW1, OW2, OW3) to extinguish a chemical fire, while also simultaneously using a single worker (OW3) to extinguish another fire. The boat captain is present during this engagement, but not depicted in this screenshot.

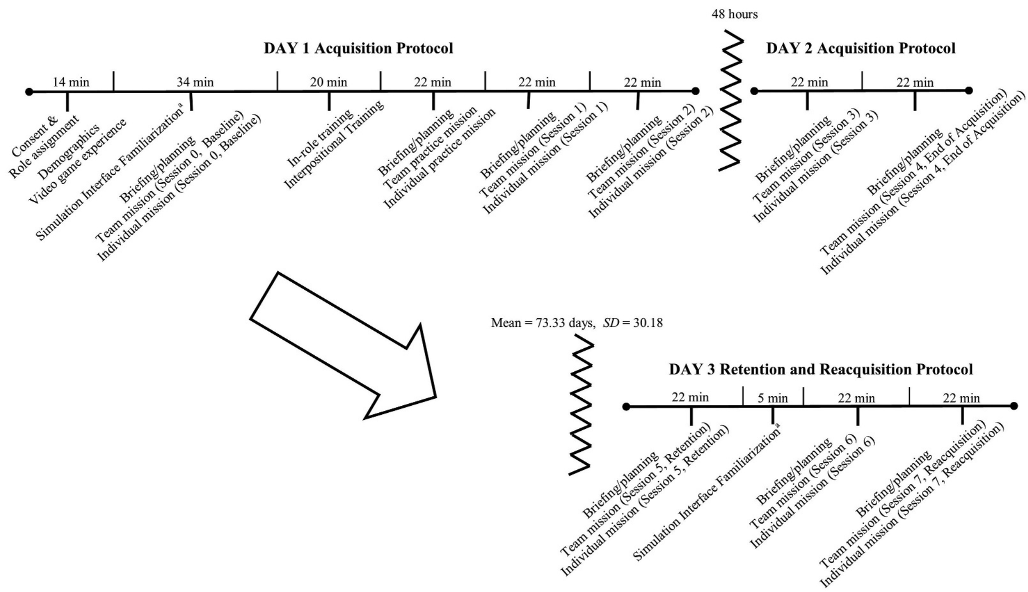

Figure 3. Overview of data collection procedure.

Note. Individual and team missions were counterbalanced for order effects. Missions lasted a maximum of 10 minutes, preceded by a two-minute briefing/planning period.[a]Participants were provided with an overview and familiarization with the simulation's interface and goals.

controlled all nine platforms simultaneously. Otherwise, the individual and team missions were identical.

Points were earned for survivors healed (10 points per survivor stabilized), survivors rescued (10 points per healed survivor picked up), and oil valves shut off (50 points per valve shut off). Each mission had 20 survivors and four oil valves; thus, the maximum attainable score for individual or team performance was 600 points. The method used to determine performance scores was explained to participants during training and scores were available during mission performance.

Procedure

Participants self-selected into available three-person lab time slots and upon arrival at the lab they were randomly assigned to one of three team roles (oilrig workers, helicopter aviators, or boat captains). Three-person teams operated the simulator collectively for team missions and individually for individual missions over the course of the study. The study protocol consisted of three phases: an acquisition, retention, and reacquisition phase. The acquisition phase was a 2-day protocol, 2½ hours long for each day with the two days being separated by a 48-hour interval. Figure 3 presents an overview of the study and data collection procedure.

Importantly, the order of team and individual missions[3] (baseline and test missions) was counter-balanced. The rationale for counterbalancing the individual and team missions was that the learning that results from performing as an individual (or as a team) might carry over to subsequent team (or individual) mission performance, thus confounding team (or individual) performance with practice effects. This was deemed particularly relevant for interpreting retention performance scores. That is,

[3]The terms mission, performance episode, and trial are used interchangeably. However, the term "mission" as used here in the paper is specific to the context of the specific task (i.e., *Crisis* missions). In contrast, "performance episode" and "trial" are more generic and are therefore used in reference to and following from the broader skill acquisition and decay literature.

the amount of decay should be higher for the first team (or individual) mission than the amount of decay of subsequent individual (or team) missions.

It is important to note that because mission order was counterbalanced during acquisition, due to attrition more participants performed the task first as individuals and then as a team. Specifically, 17 teams performed the task first as individuals and then with their corresponding team, whereas 10 teams performed the task in reverse order. Conversely, 51 individuals performed the task first as individuals and then as a team, whereas 30 individuals performed the task in reverse order.

Acquisition

On Day 1 of the acquisition phase, participants first completed individual and team baseline missions on the performance task. Then, participants received prerecorded in-role and interpositional tutorials, which were self-guided and interactive. Following the tutorials, participants completed unscored individual and team practice missions. A task aid (see Figure 4) was available onscreen (of the second monitor) during training (both acquisition and reacquisition) and performance. Participants then completed six missions (three individual and three team missions), alternating between individual and team missions. As previously noted, the order of individual and team missions was counterbalanced such that about half of the participants completed an individual mission followed by a team mission, whereas the other half completed a team mission first and then an individual mission. Prior to all team missions, participants engaged in 2 minutes of planning with their teammates where they were encouraged to formulate a mission strategy. Day 2 of acquisition was 48 hours after Day 1, and on returning to the lab, participants completed

Crisis in the Kodiak: Oilrig Search and Rescue Training Aid

Fire	What is needed to put out the fire?
Structural Fire	Water extinguishing
Electrical Fire	Chemical extinguishing
Oil Fire	Water extinguishing and Chemical extinguishing

Injury	What is needed to heal the injury?
Burn	Bandages and oxygen
Hypothermia	Blankets and oxygen
Hemorrhage	Blankets and bandages

Platform	Available First Aid Items	Available Fire Extinguishing Capabilities
Oilrig Worker	Bandages Oxygen	Water extinguishing Chemical extinguishing
Boat	Blankets Oxygen	Water extinguishing
Helicopter	Blankets Bandages	Chemical extinguishing

Figure 4. Onscreen Task/Training aid.

Note. For all *Crisis* missions, participants engaged with the simulation on their right-hand monitor, while the task/training aid was provided to participants on their left-hand monitor. Both the *Crisis* missions and task/training aid were full-screen and thus covered the entire space of their respective (dual) monitors.

two additional individual and two team missions, alternately. Again, the order of individual and team missions was counterbalanced.

Retention and reacquisition

Participants returned to the lab to complete the retention and reacquisition phase of the study an average of 73.33 days (*SD* = 30.18 days, minimum = 40 days, maximum = 169 days, median = 61 days) after Day 2. Participants were scheduled with the goal of at least a 6-week (i.e., 42-day) nonuse interval, however, this was not fixed due to the challenges of scheduling participants between semesters. There was no maximum cut off. Like the acquisition phase, the order of individual and team missions was once again counterbalanced, maintaining the acquisition phase order such that if a team completed individual before team missions at acquisition, then they also completed individual before team missions during the retention/reacquisition phase as well. Again, as indicated in Figure 3, prior to all team missions, participants engaged in 2 minutes of planning with their teammates.

Data analysis

Data analysis followed Bliese and Lang's (2016) multilevel approach for modeling discontinuous growth models, implemented with the R package *nlme* (Version 3.1–162; Pinheiro et al., 2023). In this analytical approach, several time variables are introduced to represent linear change during acquisition, decay, and reacquisition (recovery). Although time can be coded to represent either relative or absolute change in performance, only the latter was utilized to examine the present research questions (see Bliese & Lang, 2016 for further details). Accordingly, the time coding presented in Table 1 reflects absolute change and should be interpreted as follows. First, the TIME variable represents the average increase (or decrease) in performance per additional session during acquisition—that is, the sample's average performance change from Session 0 (baseline performance) to Session 4 (end of acquisition). Second, a transition variable (TRANS) was introduced to represent performance changes between the last acquisition session (Session 4) and the retention session (Session 5). Given the particular way in which TIME was coded, the TRANS variable here represents absolute change in performance following the nonuse period. Thus, the coefficient for TRANS reflects the (raw) average score difference between Session 4 and Session 5, where the sign of the TRANS variable reflects the direction of said difference such that, for instance, a negative coefficient for TRANS indicates that performance scores *decreased* after the nonuse period. Third, the recovery variable (RECOV) represents the rate of change during reacquisition (or recovery) after the nonuse period. Finally, the squared versions of the TIME and RECOV variables (TIME.SQ and RECOV.SQ) were introduced to test for potential non-linear performance trajectories during acquisition and reacquisition.

After establishing a baseline model for the full sample, additional models were estimated to compare individual and team skill acquisition, retention (decay), and reacquisition. Specifically, potential interaction effects between performance conditions (individual vs. team) and the time

Table 1. Coding of time variables for examining acquisition, decay, and recovery (reacquisition).

Session	TIME	TRANS	RECOV	TIME.SQ	RECOV.SQ
0	0	0	0	0	0
1	1	0	0	1	0
2	2	0	0	4	0
3	3	0	0	9	0
4	4	0	0	16	0
Nonuse period					
5	4	1	0	16	0
6	4	1	1	16	1
7	4	1	2	16	4

Note. TIME = acquisition; TRANS = transition (decay); RECOV = recovery; SQ = squared.

variables were examined to determine the extent to which skill acquisition, retention (decay), and reacquisition performance varied across individual and team performance. For instance, a statistically significant interaction between condition and TRANS would demonstrate the presence of differences between individual and team performance decay. In addition to the statistical significance test for each effect, we contrasted each new model to the preceding one using log-likelihood ratios (LR) and provide adjusted effect sizes analogous to Hedges's *g* (Pustejovsky et al., 2014).

Results

Individual- and team-level performance descriptive statistics and intercorrelations are presented in Table 2. The extremely low baseline scores (and variance) are not surprising given the complexity of the task coupled with participants having no previous experience with it (correlation with video game experience = −.05, $p > .05$). To address the study's objective of comparing skill acquisition, retention (decay/loss), and reacquisition for individual and team performance on a complex task, as previously noted, the data were analyzed per Bliese and Lang's (2016) multilevel approach for modeling discontinuous growth models. The results of these analyses are presented in Table 3.

Baseline model

The purpose of Models 1 through 4 was to establish a baseline model for comparing, at a later point, the trajectory of team and individual performance across acquisition, retention, and reacquisition. Model 1 in Table 3 was estimated using only the change parameters (TIME, TRANS, and RECOV). Model 1 included a term to account for autocorrelation because adding this term improved model fit significantly in comparison to a model without it, LR = 91.30, $p < .05$. This was not surprising because adjacent performance events tend to be more (positively) correlated than more distal performance events. Model 2 included the nonlinear effects for TIME and RECOV (TIME.SQ and RECOV.SQ, respectively). Compared to Model 1, Model 2 overall fit was significantly better, LR = 18.31, $p < .05$. Finally, the subsequent model (Model 3) included the random effects for TIME, TIME.SQ, TRANS, and RECOV (but not RECOV.SQ). This model offered a significant improvement in model fit, LR = 206.51, $p < .05$, suggesting that meaningful differences in acquisition, retention, and reacquisition could be explained by including additional (Level 2) variables. Dropping random effect covariances from Model 3 improved model fit even further, LR = 90.37, $p < .05$.

As previously noted, there was some variation in the length of the nonuse interval (i.e., the number of days that elapsed between the acquisition and reacquisition sessions). To examine this effect, the nonuse interval was mean-centered and then included as a predictor in the model. Model 4 shows a negative TRANS × Nonuse-interval interaction, which suggests an

Table 2. Individual and team performance descriptive statistics and intercorrelations amongst study variables.

	Team performance							
Individual performance	*M*	*SD*	1	2	3	4	5	6
1. Video game experience	1.73	0.67	–	.18	.03	.12	.15	.03
2. Performance (Session 0)	3.83	17.65	−.05	–	.30	.47*	.28	−.21
3. Performance (Session 4)	141.48	83.25	.25*	.09	–	.44*	.32	−.35*
4. Performance (Session 5)	86.25	70.75	.29*	.09	.64*	–	.49*	−.18
5. Performance (Session 7)	162.10	80.39	.19*	.07	.67*	.66*	–	−.25
6. Nonuse interval	73.33	30.18	.00	.18	−.02	−.28*	−.23*	–
M	–	–	1.71	7.41	278.52	210.74	299.26	73.33
SD	–	–	0.37	18.31	83.42	112.69	91.52	30.57

Note. Individual performance correlations are below the diagonal, and team performance correlations are above the diagonal. Individual $N = 81$, Team $N = 27$. Nonuse interval is in days. Session 0 = baseline, Session 4 = end of acquisition, Session 5 = retention, and Session 7 = reacquisition.

*$p < .05$ (two-tailed).

Table 3. Longitudinal random coefficient growth models with task performance as the dependent variable.

Effect	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7
Predictor effects							
Intercept	16.07* (8.01)	8.14 (8.27)	10.38* (4.94)	9.67 (5.73)	6.60 (6.05)	2.98 (6.23)	3.04 (6.23)
TIME	41.48* (2.20)	62.66* (6.27)	61.58* (5.87)	61.41* (5.81)	44.50* (6.41)	44.51* (6.41)	44.51* (6.40)
TRANS	−60.45* (6.53)	−57.63* (6.59)	−55.11*(6.75)	−54.62* (6.57)	−53.66* (7.35)	−53.71* (7.33)	−53.58* (8.68)
RECOV	38.90* (4.02)	65.15* (11.24)	65.50*(10.86)	71.10* (11.19)	70.97* (11.13)	70.97* (11.14)	69.31* (11.22)
TIME SQ		−5.36* (1.47)	−5.37*(1.30)	−5.38* (1.29)	−3.07* (1.46)	−3.07* (1.47)	−3.07* (1.47)
RECOV SQ		−12.41* (5.25)	−12.58*(5.14)	−12.58* (5.20)	−12.57* (5.19)	−12.58* (5.19)	−12.58* (5.18)
Nonuse interval[a]				0.17 (0.13)	0.15 (0.11)	0.15 (0.11)	0.15 (0.11)
Trial order				2.86 (7.91)	3.69 (7.02)	13.43 (8.06)	13.26 (8.06)
TRANS × Nonuse interval[a]				−0.60* (0.17)	−0.63* (0.17)	−0.63* (0.16)	−0.64* (0.16)
RECOV × Trial order				−15.27* (6.13)	−15.03* (5.97)	−15.03* (5.96)	−10.53 (7.05)
Condition					11.81 (10.92)	26.02* (12.46)	24.97* (12.46)
TIME × Condition					67.27* (12.82)	67.32* (12.82)	67.33* (12.80)
TIME.SQ × Condition					−9.22* (2.92)	−9.23* (2.93)	−9.23* (2.93)
TRANS × Condition					−2.68 (13.40)	−2.71 (13.35)	18.63 (15.64)
Condition × Trial order						−38.54* (16.09)	−35.71* (16.11)
TRANS × Trial order							−0.28 (13.07)
TRANS × Condition × Trial order							−57.92* (23.03)
Random effects							
σ^2	4481.46	4408.03	2727.63	2634.30	2537.23	2553.53	2546.98
τ_{00}	3021.13	3028.33	14.54	107.19	12.18	6.70	10.59
τ_{11}							
TIME			506.30	490.87	268.30	251.90	247.23
TIME SQ			0.02	0.02	0.00	0.00	0.01
TRANS			846.22	551.18	578.65	541.09	402.81
RECOV			13.55	9.23	9.97	7.68	34.38
ICC			0.66	0.66	0.53	0.51	0.50
σ^2	4481.46	4408.03	2727.63	2634.30	2537.23	2553.53	2546.98
Goodness of fit							
log-Likelihood	−4821.66	−4808.65	−4750.58	−4737.57	−4685.15	−4678.66	−4667.38
df	7	9	13	17	21	22	24

Note. $N = 108$ (81 individuals, 27 teams). Standard errors shown in parentheses. TIME = dummy variable for estimating (absolute) change from baseline to end-of-acquisition performance; TRANS = dummy variable for estimating skill decay between the end-of-acquisition performance session (coded 0) and the retention session (coded 1); RECOV = dummy variable to examine the reacquisition slope; SQ = squared. Dummy codes for Condition are 0 = individual and 1 = team; and for Trial order are 0 = individual followed by team session and 1 = team followed by individual session.

[a]grand-mean centered.

*$p < .05$ (two-tailed).

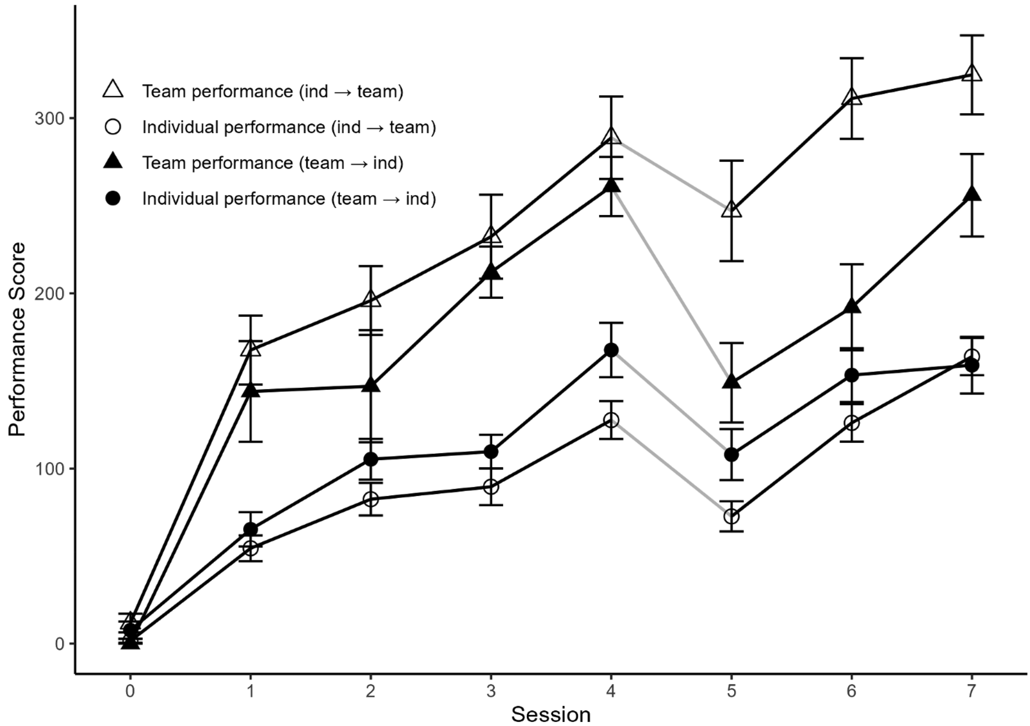

Figure 5. Individual and team performance scores during acquisition, retention (decay) and reacquisition by trial order (individual mission followed by team mission, and vice versa).

Note. $N = 108$ (81 individuals, 27 teams). Participants from 17 teams performed the task first as individuals and then with their corresponding team (individual → team) whereas 10 teams performed the task in reversed order (team → individual). Conversely, 51 individuals performed the task first as individuals and then as a team (individual → team) whereas 30 individuals performed the task in reversed order (team → individual). Nonuse period occurred between Session 4 and 5 (gray line). Error bars represent standard errors.

average decrease in performance of 0.60 points per day following the end-of-acquisition session. Model 4 also takes into account the counterbalancing of team and individual missions (i.e., trial order), and the RECOV × Trial order effect—which was the only statistically significant interaction between trial order and the change parameters. Specifically, participants from 17 teams (51 individuals) performed the task first as individuals and then with their corresponding team, whereas 10 teams (30 individuals) performed the task in reverse order. Consequently, Model 4—which included the linear and nonlinear time effects, the set of (uncorrelated) random effects mentioned previously, and the (mean-centered) nonuse interval and trial order as covariates—served as the baseline model for subsequent analyses.

Team vs. Individual acquisition

Model 5 indicated a general improvement in performance from baseline to end-of-acquisition. Specifically, performance increased by 44.50 points per trial (on average). However, as indicated by the positive TIME × Condition interaction, team performance improved faster than individual performance ($\hat{\gamma} = 67.27$, $g = 1.27$, 95% CI [0.77, 1.76]) during acquisition. In addition, the negative TIME. SQ × Condition interaction ($\hat{\gamma} = -9.22$, $g = -0.18$, 95% CI [−0.30, −0.07]) indicates that, in comparison to individual performance, the rate of acquisition of team performance slowed down faster toward the end-of-acquisition (see Figure 5).

Team vs. Individual retention and reacquisition

The TRANS × Condition term in Model 5 suggested that team and individual performance did not differ in the magnitude of decay after the nonuse period ($\hat{\gamma} = -2.68$, $g = -0.05$, 95% CI [−0.52, 0.43]). However, because Model 5 did not take into account the effect of trial order on retention, two additional models (Model 6 and 7) were estimated. Model 6 showed a statistically significant Condition × Trial order interaction. Consequently, a three-way interaction (TRANS × Condition × Trial order) was introduced in Model 7 to determine the magnitude of decay in task performance in tandem with the trial order effect. The significant effect of the TRANS × Condition × Trial Order interaction ($\hat{\gamma} = -57.92$, $g = -1.07$, 95% CI [−1.91, −0.21]) suggests that trial order affected retention performance (Figure 5). A simple effects analysis indicated that the difference between end-of-acquisition performance and retention performance for teams whose members first performed the task as a team ($\Delta_{t5-t4} = 114.21–207.36 = -93.15$) and those who first performed it as individuals ($\Delta_{t5-t4} = 198.81–233.76 = -34.95$) was significant, $\hat{\gamma} = -58.20$, $g = -1.07$, 95% CI [−1.84, −0.29]. In contrast, performance decay for individual performance was about the same regardless of whether the task was first performed as individuals ($\Delta_{t5-t4} = 60.07–113.65 = -53.58$) or as a team ($\Delta_{t5-t4} = 69.10–122.96 = -53.86$) during retention, $\hat{\gamma} = -0.28$, $g = -0.01$, 95% CI [−0.48, 0.47].

The significant reacquisition parameters (RECOV and RECOV.SQ) in Models 1–4 indicate a positive improvement in performance after the retention session with a slight deceleration from Session 6 to 7. Differences between team and individual performance during reacquisition were examined by comparing Model 5 to a model including a RECOV × Condition effect. Adding this parameter did not improve model fit, LR = 0.69, $p > .05$. For completeness, a RECOV × Condition × Trial order three-way interaction was added to Model 7 but the resultant model offered no improvement in fit either, LR = 5.09, $p > .05$. In fact, as can be seen in Table 3, the RECOV × Trial order interaction was not statistically significant once the TRANS × Condition × Trial order was added to Model 7.

In summary, the totality of the results indicates that, as expected, team performance on this complex task was substantially higher than individual performance in all phases of performance (i.e., acquisition, retention, and reacquisition). However, as illustrated in Figure 5, the amount of loss/decay was a function of whether the *first* retention mission was performed as individuals or as a team. Specifically, team performance when the first retention mission was performed as a team displayed the greatest skill decay.[4] Finally, although team reacquisition performance improved at a slightly faster rate than that for individual performance, the differences were not statistically significant.

Discussion

As previously noted, there has been a very limited amount of team skill decay research, let alone a comparison of team versus individual skill decay. Whereas team performance improved faster during acquisition and reacquisition, although the latter was not statistically significant, a noteworthy finding of the present study is that team performance after the nonuse interval (i.e., decay) was contingent on the nature of the first retention mission. Specifically, team performance loss was the largest when participants performed the *first* retention mission as teams. In contrast, decay for individual performance was similar regardless of whether the first retention mission was first

[4]A reviewer raised the possibility that this finding was confounded by differences in the length of the nonuse interval across trial order conditions. In response to the reviewer's comment, we ran some additional post hoc analyses to explore potential differences in the nonuse interval for each condition, the findings from which are presented in the online supplementary materials (Tables S1 and S2 and Figures S1 and S2). A summary conclusion based on the findings from these analyses is that there appears to be no indication that the trial order effects are attributable to differences in the length of the nonuse interval; overall, the length of the nonuse interval for the individual → team ($M = 74.39$ days, $SD = 34.01$) and team → individual ($M = 71.70$ days, $SD = 22.18$) trial order conditions is comparable ($d = 0.09$).

performed as individuals or as a team. Thus, for individual performance loss, the order of the first retention mission did not make a difference.

The abovementioned results are congruent with the interdependent nature of the team task. When participants performed the first retention session as teams, they performed the tasks associated with *only one* of the three available roles (i.e., oilrig worker, helicopter aviator, or boat captain). Not surprisingly, their subsequent individual performance—that involved performing *all* the three roles simultaneously—did not seem to have benefited from performing the preceding session as a team. In contrast, team performance was enhanced when participants performed the first retention session as individuals. These participants reacquired skills and knowledge germane to their role and the roles of their teammates (Arthur & Day, 2013a). These results are consistent with Arthur and Day's observation that complex tasks provide trainees with increased opportunity to relearn the task during the retention phase because of the exposure to the various task parameters and components. Thus, from an applied perspective, the results of the present study illustrate how a relatively brief individual post-retention (nonuse) interval training session can mitigate decay (Arthur & Day, 2020), reducing the time it takes for a team to regain proficiency after a relatively long period of nonuse. They also add to the limited work—albeit individual-focused—on techniques to attenuate skill decay (e.g., Frank & Kluge, 2018, 2019; Klostermann et al., 2022).

The serendipitous findings for team performance wherein the trial-order effect was observed after the nonuse interval but not acquisition, further highlight the importance of recognizing that although related, acquisition, retention, and reacquisition are distinct phenomena that in their totality can provide insightful information about training (and education) programs. Indeed, as previously noted, in recognition of a limitation of the tendency of the learning, training, and educational literatures to study learning (i.e., immediate posttraining performance or amount of skill acquired) and retention *independently* (Arthur et al., 1998, 2010; Schmidt & Bjork, 1992), Schmidt and Bjork (1992) emphasized that because acquisition (i.e., amount of skill acquired) and retention and transfer are indeed separate phenomena, they may yield different interpretations of the effectiveness of training. To fully understand the effects of a training condition or manipulation, one must measure its effects in not only the acquisition phase but also the retention and transfer phases as well, as was the case here. Therefore, a contribution of the present study is the assessment of complex skill performance in the context of immediate posttraining performance (i.e., amount of skill acquired) and retention after an extended period of nonuse (i.e., long-term retention).

From a research design perspective, the present study utilized a within-subject design wherein participants performed the same task concurrently as individuals and as a team. That said, potential order and carry over effects were addressed by counterbalancing the individual and team trials, resulting in the serendipitous decay effects observed here. In addition, the nonuse interval of the present study was on average about 73 days, a relatively long nonuse interval in the context of academic research. For instance, in Wang et al.'s (2013) skill decay meta-analysis that was based on 111 data points (k), only three had nonuse intervals greater than 90 days and most ($k = 48$) were between 1–7 days. It is also noteworthy that the analytic approach for modeling growth used here provided insights into the trajectory of skill acquisition, retention (decay), and reacquisition for team and individual performance while accounting for a number of factors (e.g., nonuse interval, trial order) that would not have been gained with more traditional analytical approaches (e.g., repeated-measures ANOVA).

Consistent with Arthur and Day (2020), decay was operationalized as the difference between immediate post-training and delayed posttest scores. Consonant with this operationalization, a comparison between team and individual skill decay does not reveal, in and of itself, anything in particular about the processes or mechanisms underlying skill decay. That said, the results suggest that if the task permits the performance of the *full* task (or approximations of it) by individuals on the first post-nonuse retention performance episode, then team skill decay will be minimized. This is consistent with Arthur and Day's (2013a) observation that performing a complex task after a period of nonuse (even if it is relatively brief) has the unintended effect of mitigating the amount of skill loss by facilitating rapid reacquisition. Specifically, they note that complex tasks are characterized by more task parameters and to allow all of these parameters to run their course, longer tests of performance are required. Consequently, in the context of retention, these longer tests then naturally provide individuals with the opportunity to reacquaint themselves with the task as they test.

It is also noteworthy that the differences between team and individual performance were quite large. This is consistent with previous research (e.g., Arthur et al., 1997; Day et al., 2005; Shebilske et al., 1992) which has clearly demonstrated that teams consistently outperform individuals. A common feature of the present task (*Crisis*) and other similar synthetic tasks used in teams research is that in the individual condition, participants control all elements of the task (e.g., oilrig workers, helicopter aviators, and boat captains [3 each] in *Crisis*; and pilot-gunner, and mine-missile manager in *Space Fortress* (see Arthur et al., 1997; Day et al., 2005; Shebilske et al., 1992). In contrast, in the team condition, each of these roles is controlled by one person, thus, *Crisis* engenders 3-person teams and *Space Fortress* 2-person teams. Consequently, it would not be inaccurate to conclude that for teams, task difficulty for any given team member will be lower compared to that experienced by an individual controlling all components, and therefore the reason why teams outperform individuals on these tasks. However, because the amount of acquisition is one of the most robust predictors of skill loss (individuals/teams who have acquired more skill have more to lose; hence, if one has not acquired the skill, then one cannot lose it [Arthur & Day, 2020]), examinations of loss have to control for acquisition. In doing so, the rank order of four conditions illustrated in Figure 5 indicated that the amount of skill loss for team performance where the members performed the first retention episode as individuals, displayed a level of skill loss (34-point drop) that was less but similar to that for the two individual performance conditions (53-point drop). This is in contrast to team performance when the first retention performance episode was performed as a team (93-point drop). Thus, the results are consistent with the observation that for performance on most tasks, teams (compared to individuals) have more to lose (Arthur & Day, 2020).

Limitations and suggestions for future research

There are obvious potential limitations of the present study that warrant acknowledgment. First, although the post hoc power analysis indicated we had sufficient power, the number of teams was nevertheless relatively small and by virtue of being a lab study, in spite of having a relatively long nonuse interval compared to the extant literature, it used ad hoc teams whose members were young college undergraduates, recruited from the research participant pool of a psychology department, and also did not have any interactions with each other outside the lab, a situation that does not characterize the typical operational team. Hence, these are important boundary conditions in generalizing the findings to operational teams. The use of a synthetic task environment also warrants some discussion. As reflected in Sitzmann (2011), although computer-based simulations of this sort have clear research informative value and utility (high internal validity), they could also be described as "contrived" (low ecological validity). However, it is recognized that "STEs [synthetic task environments] provide a valuable compromise between the complexity of the real world, which is an important influence on team [and individual] performance and critical for establishing externally valid results, and experimental control, which is necessary to establish internally valid results." (Salas et al., 2008a, p. 543), which is reflected in their widespread use (Sitzmann, 2011).

A primary recommendation arising from our findings is that where possible, the first retention test for teams should be performed individually. However, although this finding has a strong conceptually basis (e.g., conditions of practice, where highly complex and organized tasks are best trained with whole instead of part training), it was nevertheless serendipitous and thus needs to be replicated on an a priori basis. The efficacy of this recommendation is also dependent on whether the operational team task permits performance by an individual. However, even if it does not, it highlights the pivotal role that cross-training could play in mitigating loss and enhancing reacquisition. If trainees cannot perform the whole task as individuals, then maybe opportunities could be provided to perform the various components of the task (e.g., Shebilske and colleagues active interlocked modeling training protocol [Arthur et al., 1997; Shebilske et al., 1992]). For instance, Volpe et al. (1996) showed that interpositional clarification—instead of a full-fledged interpositional training—was sufficient to enhance teamwork and team performance in the context of a relatively low interdependent task. Likewise, Marks et al. (2002) found that positional modeling was just as effective at helping teams develop shared mental models than cross training that involved direct hands-on experience in other roles.

The potential advantageous effect of performing the first retention session for teams individually, may also carry over into settings where the team roles during retention are the same but team membership, either via turnover, substitution, or switching, is different (Gorman & Cooke, 2011; Schulte et al., 2022). The effects of team member substitution (switching, mixing, or turnover) has been discussed more so in in terms of its deleterious effect on teamwork variables and processes, and less so on taskwork (Argote et al., 2018; Gorman & Cooke, 2011; Summers et al., 2012; van der Vegt et al., 2010). Thus, giving individuals (in teams) the opportunity to rapidly reacquire lost taskwork skills by performing the first retention session as individuals could serve as a means to mitigate the overall loss of performance (resulting primarily from teamwork loss) since taskwork loss is one less thing with which they have to be concerned. Of course, these propositions need to be empirically examined.

Finally, the present study focused exclusively on taskwork and thus, does not speak to teamwork and other process variables (Cooke et al., 2013). Specifically, as previous noted, Cooke et al. examined the retention of coordination skills. One could also examine additional team process variables such as communication, team mental models, and even team efficacy. However, to the extent that these variables require some interaction with team members, conceptually, one would not expect performance on the first retention session as an individual to have much of, if any effect on these variables. Nevertheless, future research might still seek to examine the generalizability of this effect to teamwork and other team process variables. In conclusion, recognizing the obvious need to replicate our serendipitous finding, as observed by a reviewer, our results indicate that a little individual practice can go a long way toward maintaining team retention performance, but in contrast, a little team practice does not do the same for individual retention performance.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This research was sponsored by contracts and grants from the U.S. Air Force Airman Systems Directorate, Warfighter Interactions and Readiness Division, Wright Patterson AFB, OH, and also partially funded by a Society for Military Psychology (Division 19), American Psychological Association Research Grant. The views expressed in this article are those of the authors and do not reflect the official policy or position of the Department of the Army, DOD, or the U.S. Government.

Data availability statement

Data and analysis code are available upon request. Data were analyzed using R version 4.3.1 (R Core Team, 2023) and the packages *nlme* version 3.1–162 (Pinheiro et al., 2023) and *ggplot2* version 3.3.2 (Wickham, 2016). The study design and analyses were not preregistered.

References

Argote, L., Aven, B. L., & Kush, J. (2018). The effects of communication networks and turnover on transactive memory and group performance. *Organization Science*, *29*(2), 191–206. https://doi.org/10.1287/orsc.2017.1176

Arthur, W., Jr., Bennett, W., Jr., Stanush, P. L., & McNelly, T. L. (1998). Factors that influence skill decay and retention: A quantitative review and analysis. *Human Performance*, *11*(1), 57–101. https://doi.org/10.1207/s15327043hup1101_3

Arthur, W., Jr., & Day, E. A. (2013a). A look from "aFarr" (1987). The past, present, and future of applied skill decay research. In W. Arthur, Jr., E. A. Day, W. Bennett Jr., & A. M. Portrey (Eds.), *Individual and team skill decay: The science and implications for practice* (pp. 321–343). Routledge.

Arthur, W., Jr., & Day, E. A. (2013b). Introduction: Knowledge and skill decay in applied research. In W. Arthur, Jr., E. A. Day, W. Bennett, Jr., & A. M., Portrey (Eds.), *Individual and team skill decay: The science and implications for practice* (pp. 3–13). Routledge.

Arthur, W., Jr., & Day, E. A. (2020). Skill decay: The science and practice of mitigating skill loss and enhancing retention. In P. Ward, J. M. Schraagen, J. Gore, & E. Roth (Eds.), *The Oxford handbook of expertise: Research & application* (pp. 1085–1108). Oxford University Press.

Arthur, W., Jr., Day, E. A., Bennett, W., Jr., McNelly, T. L., & Jordan, J. A. (1997). Dyadic versus individual training protocols: Loss and reacquisition of a complex skill. *Journal of Applied Psychology*, *82*(2), 783–791. https://doi.org/10.1037/0021-9010.82.5.783

Arthur, W., Jr., Day, E. A., Villado, A. J., Boatman, P. R., Kowollik, V., Bennett, W., Jr., & Bhupatkar, A. (2010). The effect of distributed practice on immediate post-training, and long-term performance on a complex command-and-control task. *Human Performance*, *23*(5), 428–445. https://doi.org/10.1080/08959285.2010.515277

Arthur, W., Jr., Day, E. A., Villado, A. J., Glaze, R. M., Schuelke, M. J., Boatman, P. R., Kowollik, V., Wang, X., & Bennett, W., Jr. (2013). A comparative investigation of individual and team skill retention and transfer on a complex command-and-control simulation task. In W. Arthur,Jr., E. A. Day, W. Bennett Jr., & A. M. Portrey (Eds.), *Individual and team skill decay: The science and implications for practice* (pp. 321–343). Routledge.

Arthur, W., Jr., Naber, A. M., Jarrett, S. M., Glaze, R. M., Schurig, I., McDonald, J. N., & Muñoz, G. J. (2011). *Game-based assessment for team training and empirical research: The development of a ddd-based research tool*. Contract No. FA8650-05-D-6502, L-3 communications corporation, link training and simulations Systems, Mesa, AZ; and department of defense, USAF air force research laboratory (AFRL), warfighter training research Division, Mesa, AZ.

Bjork, E. L., & Bjork, R. A. (2011). Making things hard on your-self, but in a good way: Creating desirable difficulties to enhance learning. In M. A. Gernsbacher, R. W. Pew, L. M. Hough, & J. R. Pomerantz (Eds.), *Psychology and the real world: Essays illustrating fundamental contributions to society* (pp. 56–64). Worth Publishers.

Bjork, R. A. (1994). Memory and metamemory considerations in the training of human beings. In J. Metcalfe, & A. Shimamura (Eds.), *Metacognition: Knowing about knowing* (pp. 185–205). MIT Press.

Bliese, P. D., & Lang, J. W. B. (2016). Understanding relative and absolute change in discontinuous growth models: Coding alternatives and implications for hypothesis testing. *Organizational Research Methods*, *19*(4), 562–592. https://doi.org/10.1177/1094428116633502

Cepeda, N. J., Vul, E., Rohrer, D., Wixted, J. T., & Pashler, H. (2008). Spacing effects in learning: A temporal ridgeline of optimal retention. *Psychological Science*, *19*(11), 1095–1102. https://doi.org/10.1111/j.1467-9280.2008.02209.x

Cooke, N. J., Gorman, J. C., Duran, J., Myers, C. W., & Andrews, D. (2013). Retention of team coordination skill. In W. Arthur, Jr., E. A. Day, W. Bennett, Jr., & A. M. Portrey (Eds.), *Individual and team skill decay: The science and implications for practice* (pp. 321–343). Routledge.

Cooke, N. J., Salas, E., Kiekel, P. A., & Bell, B. (2004). Advances in measuring team cognition. In E. Salas & S. M. Fiore (Eds.), *Team cognition* (pp. 83–106). American Psychological Association.

Cooke, N. J., & Shope, S. M. (2005). Synthetic task environments for teams: CERTT's UAV-STE. In *Handbook of human factors and ergonomics methods* (pp. 46–1–46–6). CLC Press, LLC.

Day, E. A., Arthur, W., Jr., Bell, S. T., Edwards, B. D., Bennett, W., Jr., & Mendoza, J. L., Tubre, T. C. (2005). Ability-based pairing strategies in the team-based training of a complex skill: Does the intelligence of your training partner matter? *Intelligence*, *33*(1), 39–65. https://doi.org/10.1016/j.intell.2004.09.002

Day, E. A., Arthur, W., Jr., & Gettman, D. (2001). Knowledge structures and the acquisition of a complex skill. *Journal of Applied Psychology*, *86*(5), 1022–1033. https://doi.org/10.1037/0021-9010.86.5.1022

DeChurch, L. A., & Mesmer-Magnus, J. R. (2010a). Measuring shared team mental models: A meta-analysis. *Group Dynamics: Theory, Research & Practice*, *14*(1), 1–14. https://doi.org/10.1037/a0017455

DeChurch, L. A., & Mesmer-Magnus, J. R. (2010b). The cognitive underpinnings of effective teamwork: A meta-analysis. *Journal of Applied Psychology*, *95*(1), 32–53. https://doi.org/10.1037/a0017328

Epic Games, Inc. (2004). *Unreal engine [software]* https://www.unrealengine.com/

eSim games. (2007). *Steel beasts pro P.E. Software*. http://www.steelbeasts.com/

Frank, B., & Kluge, A. (2018). Complex cognitive skill retention: The roles of general mental ability and refresher interventions in a simulated vocational setting. *Journal of Computer Assisted Learning*, *34*(5), 471–481. https://doi.org/10.1111/jcal.12251

Frank, B., & Kluge, A. (2019). Is there one best way to support skill retention? Putting practice, testing and symbolic rehearsal at a test. *Zeitschrift Für Arbeitswissenschaft*, *73*(2), 214–228. https://doi.org/10.1007/s41449-018-00136-9
Gorman, J. C., & Cooke, N. J. (2011). Changes in team cognition after a retention interval: The benefits of mixing it up. *Journal of Experimental Psychology: Applied*, *17*(4), 303–319. https://doi.org/10.1037/a0025149
Hughes, A. M., Gregory, M. E., Joseph, D. L., Sonesh, S. C., Marlow, S. L., Lacerenza, C. N., King, H. B., Salas, E., & Salas, E. (2016). Saving lives: A meta-analysis of team training in healthcare. *Journal of Applied Psychology*, *101*(9), 1266–1304. https://doi.org/10.1037/apl0000120
Hughes, M. G., Day, E. A., Wang, X., Schuelke, M. J., Arsenault, M. L., Harkrider, L. N., & Cooper, O. D. (2013). Learner-controlled practice difficulty in the training of a complex task: Cognitive and motivational mechanisms. *Journal of Applied Psychology*, *98*(1), 80–98. https://doi.org/10.1037/a0029821
Jarrett, S. M., Glaze, R. M., Schurig, I., Muñoz, G. J., Naber, A. M., McDonald, J. N., Bennett, W., Jr., & Arthur, W., Jr. (2016). The comparative effectiveness of distributed and collocated team after-action reviews. *Human Performance*, *29*(5), 408–427. https://doi.org/10.1080/08959285.2016.1208662
Keiser, N. L., & Arthur, W., Jr. (2021). A meta-analysis of the effectiveness of the after-action review (or debrief) and factors that influence to its effectiveness. *Journal of Applied Psychology*, *106*(7), 1007–1032. https://doi.org/10.1037/apl0000821
Keiser, N. L., & Arthur, W., Jr. (2022). A meta-analysis of task and training characteristics that contribute to or attenuate the effectiveness of the after-action review (or debrief). *Journal of Business & Psychology*, *37*(5), 953–976. https://doi.org/10.1007/s10869-021-09784-x
Klostermann, M., Conein, S., Felkl, T., & Kluge, A. (2022). Factors influencing attenuating skill decay in high-risk industries: A scoping review. *Safety*, *8*(2), 22. https://doi.org/10.3390/safety8020022
Kluge, A., & Frank, B. (2014). Counteracting skill decay: Four refresher interventions and their effect on skill and knowledge retention in a simulated process control task. *Ergonomics*, *57*(2), 175–190. https://doi.org/10.1080/00140139.2013.869357
Kozlowski, S. W., & Ilgen, D. R. (2006). Enhancing the effectiveness of work groups and teams. *Psychological Science in the Public Interest*, *7*(3), 77–124. https://doi.org/10.1111/j.1529-1006.2006.00030.x
Mane, A., & Donchin, E. (1989). The space fortress game. *Acta Psychologica*, *71*(1–3), 17–22. https://doi.org/10.1016/0001-6918(89)90003-6
Marks, M. A., Sabella, M. J., Burke, C. S., & Zaccaro S. J. (2002). The impact of cross-training on team effectiveness. *Journal of Applied Psychology*, *87*(1), 3–13. https://doi.org/10.1037/0021-9010.87.1.3
May, G. L., & Kahnweiler, W. M. (2000). The effect of a mastery practice design on learning and transfer in behavior modeling training. *Personnel Psychology*, *53*(2), 353–373. https://doi.org/10.1111/j.1744-6570.2000.tb00205.x
Mohammed, S., Ferzandi, L., & Hamilton, K. (2010). Metaphor no more: A 15-year review of the team mental model construct. *Journal of Management*, *36*(4), 876–910. https://doi.org/10.1177/0149206309356804
Muñoz, G. J., Cortéz, D. A., Álvarez, C. B., Raggio, J. A., Concha, A., Rojas, F. I., Arthur, W., Jr., Fischer, B. M., & Rodríguez, S. (2022). After-action reviews and long-term performance: An experimental examination in the context of an emergency simulation. *Human Factors*, *64*(4), 760–778. https://doi.org/10.1177/0018720820958848
Pinheiro, J., Bates, D. R Core Team. (2023). *Nlme: Linear and nonlinear mixed effects Models. R package version 3.1-163*. https://CRAN.R-project.org/package=nlme
Pustejovsky, J. E., Hedges, L. V., & Shadish, W. R. (2014). Design-comparable effect sizes in multiple baseline designs: A general modeling framework. *Journal of Educational and Behavioral Statistics*, *39*(5), 368–393. https://doi.org/10.3102/1076998614547577
R Core Team. (2023). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/
Salas, E., Cooke, N. J., & Rosen, M. A. (2008a). On teams, teamwork, and team performance: Discoveries and developments. *Human Factors: The Journal of the Human Factors & Ergonomics Society*, *50*(3), 540–547. https://doi.org/10.1518/001872008X288457
Salas, E., DiazGranados, D., Klein, C., Burke, C. S., Stagl, K. C., Goodwin, G. F., & Halpin, S. M. (2008b). Does team training improve team performance? A meta-analysis. *Human Factors: The Journal of the Human Factors & Ergonomics Society*, *50*(6), 903–933. https://doi.org/10.1518/001872008X375009
Sauer, J., Hockey, G. R. J., & Wastell, D. G. (2000). Effects of training on short-and long-term skill retention in a complex multiple-task environment. *Ergonomics*, *43*(12), 2043–2064. https://doi.org/10.1080/00140130010000893
Schmidt, R. A., & Bjork, R. A. (1992). New conceptualizations of practice: Common principles in three paradigms suggest new concepts for training. *Psychological Science*, *3*(4), 207–218. https://doi.org/10.1111/j.1467-9280.1992.tb00029.x
Schuelke, M. J., Day, E. A., McEntire, L. E., Boatman, P. R., Boatman, J. E., Kowollik, V., & Wang, X. (2009). Relating indices of knowledge structure coherence and accuracy to skill-based performance: Is there utility in using a combination of indices? *Journal of Applied Psychology*, *94*(4), 1076–1085. https://doi.org/10.1037/a0015113
Schulte, B. D., Arthur, W., Jr., Traylor, Z., George, F., Jr. Keiser, N. L., & Bennett, W., Jr. (2022). *Effects of team member substitution and the after-action review on team performance*. In *Paper presented at the 37th Annual Conference of the Society for Industrial and Organizational Psychology*, Seattle, WA.

Shebilske, W. L., Goettl, B. P., Corrington, K., & Day, E. A. (1999). Interlesson spacing and task-related processing during complex skill acquisition. *Journal of Experimental Psychology: Applied*, *5*(4), 413–437. https://doi.org/10.1037//1076-898X.5.4.413

Shebilske, W. L., Regian, W. J., Arthur, W., Jr., & Jordan, J. A. (1992). A dyadic protocol for training complex skills. *Human Factors*, *34*(3), 369–374. https://doi.org/10.1177/001872089203400309

Sitzmann, T. (2011). A meta-analytic examination of the instrumental effectiveness of computer-based simulation games. *Personnel Psychology*, *64*(2), 489–528. https://doi.org/10.1111/j.1744-6570.2011.01190.x

Sobel, H. S., Cepeda, N. J., & Kapler, I. V. (2011). Spacing effects in real-world classroom vocabulary learning. *Applied Cognitive Psychology*, *25*(5), 763–767. https://doi.org/10.1002/acp.1747

Summers, J. K., Humphrey, S. E., & Ferris, G. R. (2012). Team member change, flux in coordination, and performance: Effects of strategic core roles, information transfer, and cognitive ability. *Academy of Management Journal*, *55*(2), 314–338. https://doi.org/10.5465/amj.2010.0175

van der Vegt, G. S., Bunderson, S., & Kuipers, B. (2010). Why turnover matters in self-managing work teams: Learning, social integration, and task flexibility. *Journal of Management*, *36*(5), 1168–1191. https://doi.org/10.1177/0149206309344117

Villado, A. J., & Arthur, W., Jr. (2013). The comparative effect of subjective and objective after-action reviews on team performance on a complex task. *Journal of Applied Psychology*, *98*(3), 514–528. https://doi.org/10.1037/a0031510

Volpe, C. E., Cannon-Bowers, J. A., Salas, E., & Spector, P. E. (1996). The impact of cross-training on team functioning: An empirical investigation. *Human Factors: The Journal of the Human Factors & Ergonomics Society*, *38*(1), 87–100. https://doi.org/10.1518/001872096778940741

Wang, X., Day, E. A., Kowollik, V., Schuelke, M. J., & Hughes, M. G. (2013). Factors influencing knowledge and skill decay after training: A meta-analysis. In W. Arthur, Jr., E. A. Day, W. Bennett, Jr., & A. M. Portrey (Eds.), *Individual and team skill decay: The science and implications for practice* (pp. 68–116). Routledge.

Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis*. Springer-Verlag. https://ggplot2.tidyverse.org

Woollard, M., Whitfield, R., Newcombe, R. G., Colquhoun, M., Vetter, N., & Chamberlain, D. (2006). Optimal refresher training intervals for AED and CPR skills: A randomised controlled trial. *Resuscitation*, *71*(2), 237–247. https://doi.org/10.1016/j.resuscitation.2006.04.005