Received: 18 March 2020 | Revised: 27 May 2020 | Accepted: 27 May 2020 | Published online: 8 July 2020

DOI: 10.1002/dad2.12055

Alzheimer's & Dementia
Diagnosis, Assessment & Disease Monitoring

RESEARCH ARTICLE

Lower practice effects as a marker of cognitive performance and dementia risk: A literature review

Roos J. Jutten[1] | **Evan Grandoit**[2] | **Nancy S. Foldi**[3] | **Sietske A. M. Sikkes**[1] | **Richard N. Jones**[4] | **Seo-Eun Choi**[5] | **Melissa L. Lamar**[6] | **Diana K. N. Louden**[7] | **Joanne Rich**[7] | **Douglas Tommet**[4] | **Paul K. Crane**[5] | **Laura A. Rabin**[8]

[1] Alzheimer Center Amsterdam, Amsterdam Neuroscience, Amsterdam UMC, Vrije Universiteit, Amsterdam, the Netherlands

[2] Northwestern University, Evanston, Illinois, USA

[3] Queens College and The Graduate Center of The City University of New York, Queens, New York, USA

[4] Warren Alpert Medical School of Brown University, Providence, Rhode Island, USA

[5] School of Medicine, University of Washington, Seattle, Washington, USA

[6] Rush Alzheimer's Disease Center, Rush University Medical Center, Chicago, Illinois, USA

[7] University Libraries, University of Washington, Seattle, Washington, USA

[8] Brooklyn College and The Graduate Center of The City University of New York, Brooklyn, New York, USA

Correspondence
Roos J. Jutten, Alzheimer Center Amsterdam, Amsterdam Neuroscience, Amsterdam UMC, location VUmc, P.O. Box 7057, 1007 MB Amsterdam, the Netherlands.
Email: r.jutten@vumc.nl

Abstract

Background: Practice effects (PEs) are improvements in performance after repeated exposure to test materials, and typically viewed as a source of bias in repeated cognitive assessments. We aimed to determine whether characterizing PEs could also provide a useful marker of early cognitive decline.

Methods: We conducted a systematic review of the literature, searching PsycInfo (Ebsco) and PubMed databases for articles studying PEs in aging and dementia populations. Articles published between 1920 and 2019 were included.

Result: We identified 259 articles, of which 27 studied PEs as markers of cognitive performance. These studies consistently showed that smaller, less-robust PEs were associated with current diagnostic status and/or future cognitive decline. In addition, lower PEs were associated with Alzheimer's disease risk factors and neurodegeneration biomarkers.

Conclusion: PEs provide a potentially useful marker of cognitive decline, and could prove valuable as part of a cost-effective strategy to select individuals who are at-risk for dementia for future interventions.

KEYWORDS
Alzheimer's disease, cognition, learning effects, practice effects, retest effects

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.
© 2020 The Authors. *Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring* published by Wiley Periodicals, Inc. on behalf of Alzheimer's Association

1 | BACKGROUND

Practice effects (PEs) are expected improvements in cognitive performance seen on repeated exposure to test material in the absence of intervention.[1] PEs, also referred to as retest or learning effects, are typically viewed as a source of bias or error when analyzing data from repeated cognitive assessments,[2,3] particularly in the study of Alzheimer's disease (AD) and other neurodegenerative disorders leading to dementia where cognitive decline is a key marker of clinical change.[4,5] PEs can hinder our understanding of the disease course of AD and other neurodegenerative diseases, as well as improve the evaluation of interventions that aim to slow or halt cognitive decline.[6] That is, by masking cognitive decline due to an underlying neurodegenerative process or by inflating cognitive gain in the absence of treatment induced brain changes, PEs may lead to underestimating the severity of disease progression or overestimating the efficacy of treatment effects.[7]

The absence of PEs may also provide useful information in the context of AD and dementia. More specifically, one might expect attenuated PEs in disorders such as mild cognitive impairment (MCI) and AD, in which learning is compromised. This was, for example, supported by Duff et al.,[8] who showed that lower PEs were predictive of cognitive decline 1 year later in individuals with MCI. Another study found that cognitively healthy older adults who later progressed to AD dementia had substantially lower PEs on episodic memory tasks compared to those who remained cognitively healthy.[9] Together, these findings suggest that lower PEs may indicate a subtle cognitive impairment preceding overt reduction in cognitive performance, and may serve as an early marker to differentiate neurodegeneration from healthy cognitive aging. This would be of particular relevance in pre-dementia disease stages such as MCI or subjective cognitive decline, when objective cognitive decline is modest or not easily captured by traditional cognitive assessments.[10,11] Therefore, the aim of the current review was to examine the role of PEs as a potential marker for cognitive decline in the study of cognitive aging.

Previous summaries of the PE literature in aging populations have largely focused on PEs as a source of bias. An example is the meta-analysis by Calamia et al., which examined the magnitude of PEs on several widely-applied cognitive tests (both memory and non-memory tests), and investigated the influence of age, test-retest interval, use of alternate forms, and clinical diagnosis on those effects.[12] Of interest, they found that clinical groups (ie, patients with neurological or psychiatric conditions) showed lower PE on average compared to cognitively healthy adults. The authors concluded that PEs should be accounted for in cognitively healthy populations to accurately assess group-level changes. Moreover, this finding also suggests that lower PEs in patient groups may reflect a cognitive (ie, learning) deficit, which could serve as a clinical marker of interest. The potential value of PEs as an indicator or marker of dementia risk is further supported by other studies, which suggest that PEs could serve as a proxy of specific fluid or imaging AD biomarkers,[13] or could be used in combination with those biomarkers to identify individuals at greatest risk for clinical progression.[14]

HIGHLIGHTS

- We reviewed published research that studied practice effects (PEs) as markers of cognitive performance
- Lower PEs may associate with current cognitive status and predict future decline
- Lower PEs may associate with specific biological risk factors for Alzheimer's disease

RESEARCH IN CONTEXT

1. **Systematic review**: The authors conducted a systematic review of the literature, searching PsycInfo and PubMed databases for articles studying practice effects (PEs) on cognitive testing in aging and dementia populations. Of 259 identified articles, 27 studied PEs as a clinically useful marker of cognitive performance in older adults.
2. **Interpretation**: We found accumulating evidence that lower PEs may represent an early indicator of cognitive decline, and that lower PEs associate with specific Alzheimer's disease (AD) biological risk factors. The combination of quantifying PEs and assessing AD biomarkers may yield an optimal approach to estimate AD risk.
3. **Future directions**: Future research should focus on identifying high-risk individuals from a combination of cognitive and clinical features, AD biomarkers, and lower than expected PEs. This could yield a cost-effective strategy to enrich samples in clinical trials and provide a valuable cognitive marker for subtle pharmacological or treatment responses.

Although PEs have been investigated and addressed in many studies for over a century, we sought to consolidate and compare studies by conducting a literature review of published research relevant to aging and dementia. More specifically, we aimed to summarize research that investigated whether the presence or absence, or magnitude of PEs could (1) be an indicator of current diagnostic status (ie, cognitively normal, MCI, or dementia); (2) predict future cognitive decline or progression to dementia; and/or (3) relate to AD risk factors or biomarkers of neurodegeneration.

2 | METHODS

2.1 | Search strategy

The authors obtained published empirical studies through a systematic search of the PsycInfo (Ebsco) and PubMed databases. Based on

consultation with research medical librarians (DKNL and JR), our search strategy used the following key search terms: practice effects, learning effects, retest effects, repeat testing, serial testing, serial assessment, longitudinal testing, neuropsychological testing, cognitive change, reliability, and early detection. A full text and abstract were required for inclusion; there were no restrictions for date of publication or language. The authors also manually examined the references of relevant studies to identify additional articles. Searches were conducted on December 14, 2017, and on April 18, 2019, December 2, 2019, and January 31, 2020, to include published manuscripts of interest.

2.2 | Screening and review process

We selected articles using the following steps:

1. *Identification*. Two authors (EG and LAR) reviewed titles and abstracts to identify peer-reviewed full-text articles focusing on PEs in aging samples or important analytic/statistical issues about modeling PEs, which were selected for subsequent evaluation.
2. *Initial screening*. Research team members (NSF, PKC, LAR, MLL, RJJ, SAMS, and RNJ) evaluated full texts of articles identified in Step 1. One investigator reviewed each article and introduced it to the group for discussion to classify the article in the subsequent step.
3. *Eligibility*. Based on the full-text evaluations, articles were categorized as (A) PEs as a nuisance variable and possible solutions; (B) measuring and understanding the construct of PEs (eg, underlying mechanisms, moderators); and/or (C) PEs as a potential measure of cognitive change. Categories were not mutually exclusive, meaning that articles could be assigned to multiple categories.
4. *Inclusion in the current study*. The goal of the current article was to determine whether PEs could provide clinically useful information in the context of dementia or dementia risk, and we therefore further evaluated articles within Category C (ie, reported practice effects as a valuable measure of cognitive change). Each Category C article was reviewed by two independent investigators (from among authors NSF, PKC, LAR, MLL, RJJ, SAMS, RNJ, and EG) and assigned to at least one theme: (1) PE variability defines cases; (2) PE variability predicts outcomes; and (3) PEs associate with AD risk factors (Table 1). Articles could be assigned both primary and secondary themes. The authors also rated each article's methodological rigor as low, medium, or high. If an article was identified as having more than one theme, it was assigned a secondary theme. If two investigators disagreed on the theme(s) or quality of an article, it was reviewed by a third investigator.

2.3 | Data extraction

The following data were extracted from each included article: study population (ie, normal cognition, MCI, and/or dementia), mean age, retest interval, cognitive domains, cognitive task, biomarkers (if available), study design (ie, cross-sectional, longitudinal, or mixed), method of quantifying PEs, research findings, and conclusions. With regard to study design, for the purpose of the current article, the term "cross-sectional" was applied to studies for which the PEs (ie, second testing to define a retest effect) were measured concurrently with the outcome of interest (eg, current diagnostic status); by contrast, the term "longitudinal" was applied to studies in which retesting to identify PEs preceded the outcome being measured (eg, progression to dementia, change in cognitive test scores). All data were extracted by one author (EG) and verified by a second author (LAR). Descriptive analyses on study population, age, retest interval, cognitive domains, cognitive task, and methods of quantifying PEs were performed by synthesizing the data across all of the relevant articles and presented in one overview table. Subsequently, data were summarized by theme to investigate whether PEs were associated with (1) current diagnostic status; (2) future cognitive decline or change in diagnostic status; and (3) AD risk factors.

TABLE 1 Theme definitions and related hypotheses of articles from Category C

Theme 1	**Practice Effect Variability Defines Cases** Hypothesis: Lower practice effects in mild cognitive impairment and Alzheimer's disease/dementia could serve as additional evidence for presence of prodromal or clinical dementia conditions.
Theme 2	**Practice Effect Variability Predicts Outcomes** Hypothesis: Lower practice effects may be an important early indicator of longitudinal outcomes such as cognitive decline, change in cognitive test scores, change in diagnostic status, or incidence of MCI or dementia.
Theme 3	**Practice Effects Associate With AD Risk Factors** Hypothesis: Lower practice effects are associated with various Alzheimer's disease risk factors/biomarkers.

3 | RESULTS

Figure 1 shows the results from the screening and review processes. The search identified 259 articles published between October 1920 and December 2019. We determined that 107 of these investigated either PEs on cognitive tests in aging samples or methods on modeling PEs, and underwent full-text evaluation by research team members. Of those, 27 met eligibility for Category C and were included in the current study. We assigned six of the 27 articles to primary Theme 1 (22%), 10 to primary Theme 2 (37%), and 11 to primary Theme 3 (41%). We assigned secondary themes for five of the articles (three to Theme 3 and two to Theme 1).

Table 2 displays the key study characteristics of the 27 articles considered in this article. Studies in these reports included from n = 25 to n = 1390 participants, with mean ages ranging from 53.4 to 83 years, with an overall average of 73.6 years. Studies evaluated several different diagnostic groups, retest intervals, cognitive domains and tasks, and used different methods to calculate PEs (Table 2). Most studies

TABLE 2 Overview of the n = 27 included articles studying practice effects as a marker of cognitive performance in aging and dementia populations

Research article and reference number	Retest interval	Baseline groups	Cognitive domains	Cognitive tasks	Biomarkers	Cross- sectional/ Longitudinal/Mixed	How PEs calculated/quantified	Mean age/ and total n
Cooper et al.[15]	1-4 days	AD vs. Normal Cognition	Memory (episodic)	AMT	N/A	Longitudinal	Change score	~70 years/ n = 123
Cooper et al.[16]	1 wk	AD vs. MCI vs. Normal Cognition	Memory (episodic)	AMT	N/A	Longitudinal	Change score	~ 69 years / n = 69
Darby et al.[17]	Same day	MCI vs. Normal Cognition	Memory (episodic, working); Attention; Psycho-motor speed	Non-standardized measures	N/A	Longitudinal	Change score	n/a / n = 60
Dodge et al.[18]	Every 18 mos (> 14 yrs)	Normal Cognition	Learning; Memory; Language; Psycho-motor speed; Executive function	MMSE; TMT-A & B; CERAD; BNT	N/A	Mixed	T-test of difference score	~72 years / n = 1,230
Duff et al.[19]	1 wk	MCI vs. Normal Cognition	Visual processing speed; Psychomotor speed; Memory (verbal, visual)	HVLT-R; BVMT-R; SDMT; TMT-A & B	Hippocampal volume	Longitudinal	Z-score difference	77.5 years / n = 25
Duff et al.[20]	1 wk and 1 year	MCI vs. Normal Cognition	Memory; Non-memory	BVMT-R; HVLT-R; COWAT; Animal fluency; TMT-A & B; SDMT	N/A	Longitudinal	Change score	78.7 years / n = 127
Duff et al.[21]	2 wks, 3 mos., & 6 mos	MCI	Visual learning; Global cognitive function; Psycho-motor speed; Memory; Language	BVMT-R; MMSE; TMT-A & B; HVLT-R; WAIS-R; BNT; COWAT; SDMT	N/A	Longitudinal	Correlation/ regression	72.4 years / 3 different samples (n = 80, n = 33, n = 170)
Duff et al.[22]	1 wk	MCI vs. Normal Cognition	Memory; Visuospatial constructive skills; Language; Attention	TMT-A & B; SDMT; BVMT-R; COWAT; HVLT-R; Animal fluency	N/A	Longitudinal	Change score	~80 years / n = 121
Duff et al.[23]	Same day	MCI	Global cognitive function	HVLT-R; MMSE	N/A	Longitudinal	Change score; Correlation/ regression	73.3 years / n = 61
Duff, et al.[24]	1 wk	MCI vs. Normal Cognition	Visual memory	BVMT-R	Amyloid uptake	Longitudinal	Correlation/ regression	74.6 years / n = 25
Duff et al.[25]	1 wk	MCI vs. Normal Cognition	Memory (verbal, visual); Attention & psychomotor speed; Visual scanning/ processing speed; Premorbid intellect	HVLT-R; BVMT-R; SDMT; TMT-A & B; WRAT-4	Amyloid uptake	Longitudinal	Correlation/ regression	77.5 years / n = 27

(Continues)

TABLE 2 (Continued)

Research article and reference number	Retest interval	Baseline groups	Cognitive domains	Cognitive tasks	Biomarkers	Cross- sectional/ Longitudinal/Mixed	How PEs calculated/quantified	Mean age/ and total n
Duff et al.[13]	~1 wk	MCI vs. Normal Cognition	Learning & memory; Attention & processing speed; Executive function; Premorbid intellect; Global cognitive function	HVLT-R; BVMT-R; TMT-A & B; SDMT; TMT-A; COWAT; WRAT-4; MMSE	Brain hypo-metabolism	Longitudinal	Correlation/ regression	74.6 years / n = 25
Duff et al.[8]	1 wk	MCI vs. Normal Cognition	Memory (objective, immediate, delayed); Premorbid intellect; Executive function; Attention; Language; Visuospatial constructive skills	WRAT-3; 3MS; RBANS; BVMT-R; HVLT-R; COWAT; TMT-A & B; SDMT	N/A	Mixed	Change score	~80.6 years / n = 108
Galvin et al.[26]	Yearly (x6)	Normal Cognition	Memory (primary, working, episodic, verbal); Visuospatial constructive skills; Language	WMS; BVRT; Word fluency; WAIS; TMT-A; BNT	Amyloid deposition, Braak and Braak stage, Neurofibrillary scores, Lewy bodies, cortical infarcts & hemorrhages	Longitudinal	Correlation/ regression	80.7 years / n = 80
Hanyu et al.[27]	1 wk	Amnestic MCI	Logical memory; Language; Global cognitive function	WMS-R; Category fluency; MMSE	N/A	Longitudinal	T-test difference	~ 76 years / n = 39
Hassenstab et al.[9]	~2 mos	Normal Cognition	Memory (episodic, semantic); Executive function; Visuospatial function; Global cognitive function	FCSRT; WMS; WMS-R; WAIS; BNT; Animal Naming; TMT-A & B	APOE genotype	Longitudinal	Correlation/ regression	74.5 years / n = 263
Howieson et al.[28]	Every 6 mos (x3)	Normal Cognition	Memory; Language; Visuospatial constructive skills	WMS; Category fluency; WAIS-R	N/A	Longitudinal	Correlation/ regression	83 years / n = 156
Ihara et al.[29]	3 yrs	MCI and Normal Cognition	Memory (episodic); Global cognition; Executive function; Visuospatial function	MMSE; ADAS; WMS-R; WAIS-R; Category fluency; TMT; BNT; CDT; CCT	Amyloid uptake APOE	Longitudinal	Correlation/ regression	~68 years / n = 84
Jonaitis et al.[30]	4 yrs x1 then 2 yrs x10	Normal cognition (AD family history vs. none)	Verbal learning & recall; Executive function; Visual learning & memory	RAVLT; BVMt; WMS	AD family history of AD; APOE genotype	Longitudinal	Change score	~54 years / n = 594
Machulda et al.[31]	15 mos (x3)	Normal Cognition	Memory; Language; Visuospatial function; Attention/ Executive function	AVLT; WMS-R; BNT; Category fluency; WAIS-R; TMT-B	Hippocampal volume, brain hypometabolism; Amyloid status	Longitudinal	Z-score difference	~75 years / n = 190

(Continues)

TABLE 2 (Continued)

Research article and reference number	Retest interval	Baseline groups	Cognitive domains	Cognitive tasks	Biomarkers	Cross- sectional/ Longitudinal/Mixed	How PEs calculated/quantified	Mean age/ and total n
Machulda et al.[32]	15 mos (x2-x5)	Normal Cognition	Memory; Language; Visuospatial function; Attention/ Executive function	AVLT; WMS-R; BNT; Category fluency; WAIS-R; TMT-B	APOE genotype	Longitudinal	Z-score difference	78.1 years / n = 1,390
Oltra-Cucarella et al.[14]	6-72 mos	MCI vs. Normal Cognition	Memory	AVLT	APOE genotype	Longitudinal	Z-score difference	~74 years / n = 1,210
Sanchez-Benavides et al.[33]	6 wks	Normal Cognition (Family history of AD vs. APOE genotype)	Memory (short-term, working, visual); Processing speed; Visual perception; Coordination; Attention	WAIS-IV	APOE genotype Family history of AD	Longitudinal	T-test difference	53.43 years / n = 400
Schrijnemaekers et al.[34]	2-3 yrs	Normal Cognition	Memory Global cognitive function	HVLT; MMSE	N/A	Mixed	Correlation/ regression	~76 years / n = 101
Suchy et al.[35]	17 mos	MCI	General cognition status; Attention; Visual-constructive ability	DRS-2	N/A	Longitudinal	Correlation/ regression	~70 years / n = 75
Wilson et al.[36]	Yearly (x5)	Normal Cognition group	Memory (episodic, semantic, working); Perceptual speed; Visuo-spatial ability	WMS; Category fluency; BNT; Word Reading Test; SDMT; Judgment Line Orientation; Standard Progressive Matrices	Tangles, beta-amyloid, and hippocampal volume	Longitudinal	Change point modelschange score	78.7 years / n = 567
Zehnder et al.[37]	2.4 yrs	AD vs. Normal Cognition	Global cognitive function	CERAD-NAB (German version: included the - BNT, MMSE, Animal fluency, Word list I-III, Figures-copy, Word-list delayed recall, Word list recognition, Figures-delayed recall)CDT	N/A	Longitudinal	Z-score difference	~71.25 / n = 469
Most Frequent	1 wk (8 total)	Normal Cognition (23 total)	Visuospatial/ constructive ability (10 total)	TMT (14 total)	Higher amyloid uptake on amyloid PET scans (4 found an effect and 2 did not)	Longitudinal (24 total)	Correlation/ regression (11 total)	Total average (73.57 years / n = 272)

Abbreviations: 3MS, modified mini-mental state examination; AD, Alzheimer's disease; ADAS, Alzheimer's disease assessment scale; AVLT, auditory verbal learning test; BNT, Boston naming test; BVMT(-R), brief visuospatial memory test (revised); CCT = clock copying test; CDT = clock drawing test; CERAD, consortium to establish a registry for Alzheimer's disease; COWAT, controlled oral word association test; DRS, dementia rating scale; FCSRT, free and cued selective reminding test; HVLT, Hopkins verbal learning test; MCI, mild cognitive impairment; MMSE, mini-mental state examination; RBANS, repeatable battery for the assessment of neuropsychological status; SDMT, symbol digit modalities test; TMT, trail making test; WAIS(-R), Wechsler adult intelligence scale (revised); WMS(-R), Wechsler memory scale (revised); WRAT-3, wide range achievement test.

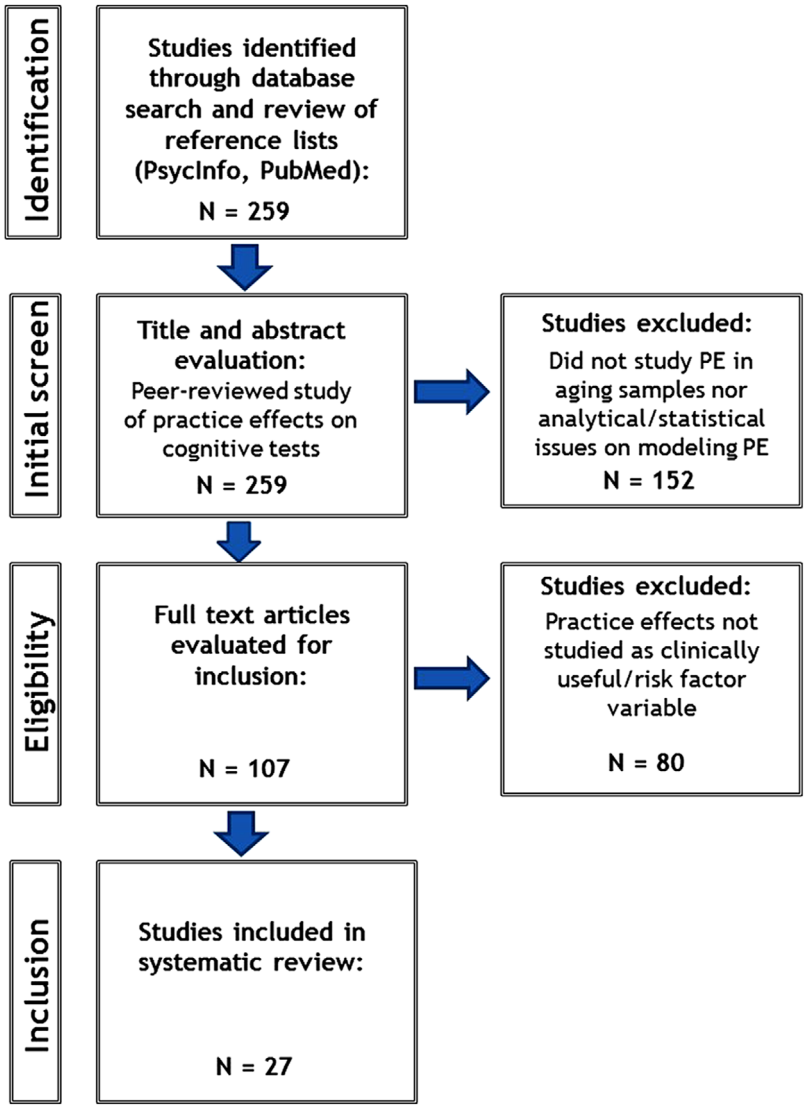

FIGURE 1 Overview of the screening and review process

(n = 23, 85%) recruited cognitively healthy older adults, whereas 15 studies (56%) recruited participants with MCI, and only 3 studies (11%) recruited individuals with AD dementia. Retest intervals ranged from same day to 4 years, where seven studies (26%) included additional multiple retest assessments (Figure 2). Ten articles included visuospatial functioning or construction, nine articles included language, nine articles included attention, eight articles included executive functioning, seven articles included episodic memory, four articles included psychosocial speed, and four articles included working memory. To calculate PEs, 12 (44%) of the studies used correlation or regression analyses, 8 (30%) used z-score or *t*-test difference scores, and 9 (33%) used a type of change score quantifier.

3.1 | Theme 1: PE variability defines cases

Studies assigned to Theme 1 addressed the hypothesis that individuals with MCI or AD had lower PEs than individuals with normal cognitive functioning.[15,16,17,22,23,37] Most of these studies showed that cognitively healthy elderly showed significantly greater PEs on average than individuals with MCI or dementia on both memory and non-memory cognitive tasks.[15,16,23,37] One study found that lower PEs in individuals with MCI could be detected after multiple repeated assessments on the same day.[17] Conversely, Duff et al. reported that some individuals with amnestic MCI had PEs in delayed recall similar to those found in cognitively healthy older adults. However, when the amnestic MCI group was split into those who remained stable ("MCI-stable," classified as MCI both at baseline and 1 week follow-up) and those who improved and subsequently appeared intact ("MCI-normal," classified as MCI at baseline but as intact at 1 week follow-up), the latter showed significant PEs, whereas the former did not.[22] Thus, it was the variability of PEs from initial to subsequent testing that served as a potential more reliable diagnostic indicator. Finally, it should be noted that results on whether PEs could differentiate clinical groups independently of baseline cognitive test performance were somewhat contradictory. For example, Zehnder et al. found that quantifying PEs did not add diagnostic accuracy to baseline cognitive test scores when discriminating cognitively healthy from AD participants.[37] In contrast,

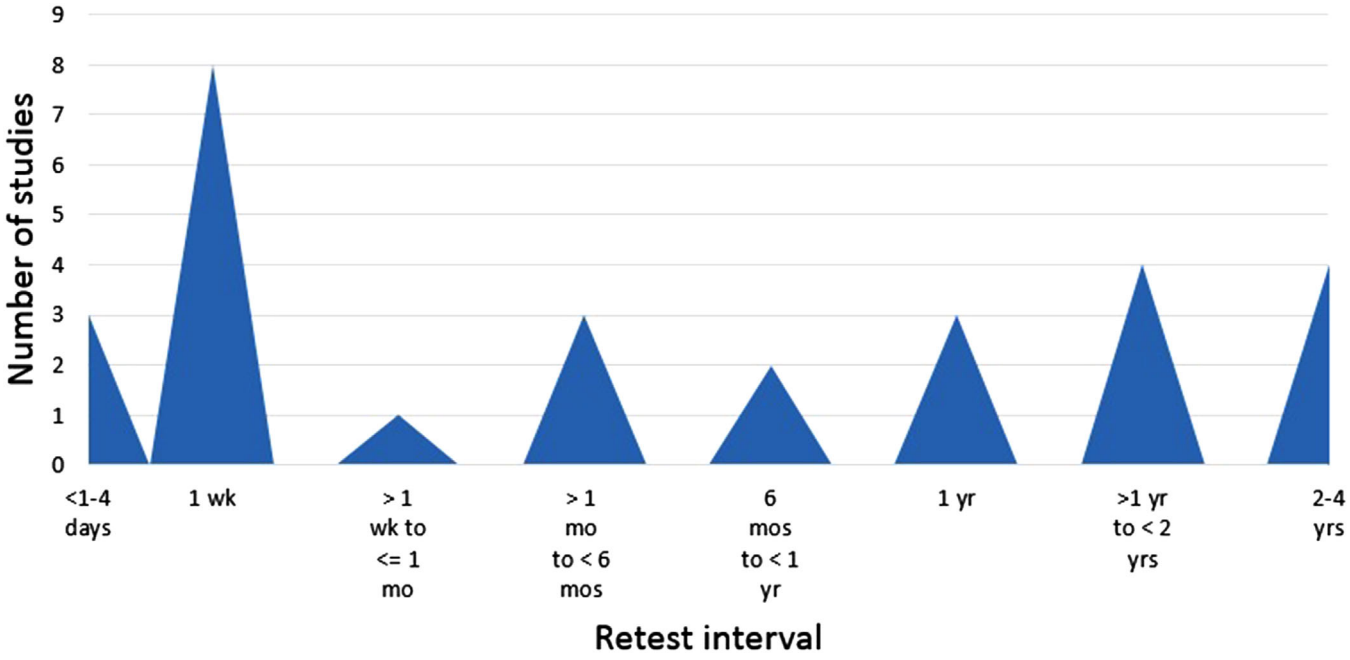

FIGURE 2 Overview of different retest intervals across all included articles. (The Oltra-Cucarella et al. (2018) article was not included due to large per subject retest interval range.)

Duff et al. showed that the predictive value of PEs was additive to baseline cognitive performance after a 2-h retest interval.[23]

3.2 | Theme 2: PE variability predicts future cognitive decline

All Theme 2 articles addressed the hypothesis that lower PEs observed over time may be an important early indicator of future cognitive decline. This was evidenced by associations between smaller or less robust PEs and (1) subsequent decline in cognitive test scores[8,20,21,35]; (2) risk of progression to AD[9]; (3) incidence of MCI or dementia[14,32,34]; and (4) terminal decline.[18] Most studies showed that findings varied by cognitive domain, with abundant evidence that lower PEs on episodic memory measures predicted future cognitive decline,[9,18,21,32,34] whereas others found that the predictive value of PEs was largely consistent across different cognitive domains.[20] Furthermore, it should be noted that studies varied highly in terms of retest intervals, which ranged from single-time retesting at the same day,[23] to multiple repeated assessments over several years[32] (Figure 2).

3.3 | Theme 3: PE variability associate with AD risk factors and biomarkers

The majority of articles we assigned to Theme 3 demonstrated that lower PEs were associated with AD risk factors and biomarkers indicative of neurodegeneration (Table 3). For example, several studies showed that lower PEs were associated with greater presence of AD risk factors such as apolipoprotein E gene (*APOE*) ε4 alleles or a positive family history of dementia.[14,32,33] In addition, attenuated PEs have been found to be more common in groups with specific neuroimaging markers that are indicative of dementia, such as lower cerebral blood flow,[27] brain hypometabolism,[13] smaller hippocampal volumes,[19] and higher levels of amyloid.[24,25,29] More specifically, Machulda et al. showed that lower PEs were associated with lower hippocampal volume and brain hypometabolism regardless of amyloidosis, suggesting that PEs were more closely related to neurodegeneration than amyloid status.[31] Studies linking PEs to neuropathology showed contrasting findings. Galvin et al. reported the relationship between PE differences and the presence of AD neuropathology among individuals who had died without a clinical diagnosis of dementia.[26] Those with AD pathology had lower PEs on episodic and semantic memory tests than did those without AD pathology. In contrast, Wilson et al. did not identify a relationship between PEs over years and post-mortem neuropathological markers of AD in individuals without dementia.[36]

4 | DISCUSSION

The literature tends to characterize PEs as a nuisance in estimating group-level characteristics such as normative decline over time in advanced age across different clinical and cognitive populations.[2,38–40] In addition, in clinical neuropsychological evaluations, neuropsychologists are typically concerned with the question of whether individuals show statistical evidence of change beyond that expected based on average PEs.[3,41–46] In the current systematic review, we focused on a different question, namely whether individual-level PEs in older adults could serve as a marker of clinical status, such that individuals

TABLE 3 Overview of studies with evidence for associations between PEs and AD risk factors or biomarkers

Finding	Number of articles showing an association	Number of articles not showing an association
Presence of ≥ 1 *APOE* ε4 allele	3 [14, 27, 32]	3 [9, 24, 36]
Higher amyloid uptake on amyloid PET scans	4 [30-32, 34]	2 [33, 35]
Lower hippocampal volume	2 [29, 33]	1 [35]
Brain hypometabolism on FDG-PET	2 [13, 33]	0
Family history	1 [36]	1 [27]
Cortical infarcts and hemorrhages	1 [34]	0
Lewy bodies	1 [34]	0
Braak stage	1 [34]	0

Abbreviations: FDG, fluorodeoxyglucose; PET, positron-emission tomography.

who showed greater PEs would be at lower risk of future negative cognitive/clinical outcomes than those who had lower PEs. Overall, we found consistent evidence from a modest-sized published literature that smaller, less robust PEs on repeated cognitive testing may be an important early indicator of current diagnostic status and future cognitive decline. In addition, lower PEs were associated with risk factors and markers indicative for dementia, such as *APOE* genotype and biological markers of neurodegeneration.

Our review identified 27 articles evaluating evidence that characterized PEs in older adults can provide a marker for (future) cognitive performance and risk of dementia. Overall, 25 of those articles provided support for one or more of our defined hypotheses (Figure 3). These studies supported the hypotheses that the magnitude of PEs is associated with (1) current cognitive performance or clinical status (Theme 1) and/or (2) future cognitive decline and risk of progression to AD dementia (Theme 2). Most of these studies were cross-sectional or longitudinal comparisons between cognitively healthy older adults and those with MCI or AD.[8,15–17] However, one study suggested that characterizing PEs in cognitively healthy individuals at baseline could aid in the prediction of who would develop AD in the future.[9] Lower PEs were also found to be more common in groups with specific biological markers (Theme 3), including brain metabolism, hippocampal volume, and amyloid load.[13,29,31] Together with Theme 2 findings, these results suggest that the assessment of PEs, in combination with biomarkers, could be used to detect preclinical AD. This would be of particular relevance in the context of current AD research and clinical trials, which are increasingly focusing on earlier, preclinical populations.[11,47]

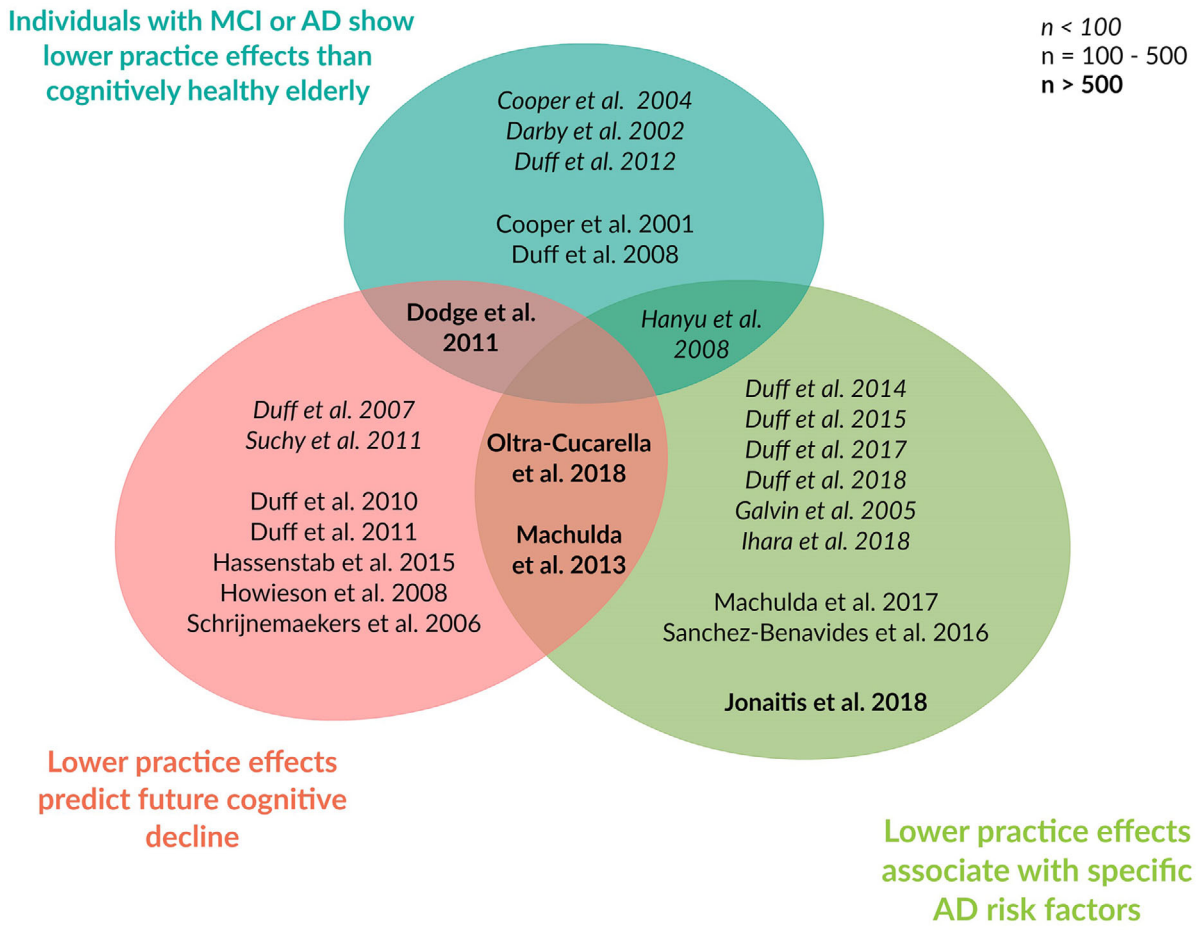

FIGURE 3 Overview of the 25 articles providing evidence for one or more of the a priori defined hypotheses (Theme 1, 2, and/or 3)

Although we identified accumulating evidence of the clinical value of PEs in the study of AD, it should be noted that some findings remain inconclusive. For instance, it is yet unclear whether PEs can predict future cognitive decline beyond baseline cognitive performance.[37] In addition, some studies showed consistent results for PEs across multiple cognitive domains,[20] whereas other studies indicated that the predictive value of PEs was domain-specific.[31] It is difficult to determine whether differences across studies reflected disparate assessment protocols or the selection of only statistically significant findings to present in publications. Findings regarding *APOE* ε4 status were also contradictory; whereas three studies found that *APOE* ε4 carriers had lower PEs for memory,[14] three other studies found no such association between *APOE* ε4 alleles and PEs.[9,30,32] Unfortunately, it was difficult to formally compare studies with contrasting findings, as the studies reported varied highly with respect to retest intervals, diagnostic groups, cognitive domains, biomarkers, and methods used to calculate and model PEs (see Table 2). Particularly the latter should be taken into account when integrating the various findings. For example, an empirical change score between the first- and second-time testing captures PEs differently than, for example, a regression-based slope of performance across multiple years. The method that reflects PEs most directly likely depends on multiple factors, such as the number and timing of the repeated assessments.[28,48] Although a full review of methods used to quantify PEs is beyond the scope of the current article, this is an important topic to address in future research.

Another important issue was the methodological complications of the included studies, such as, for example, small-sample sizes (eg, $n < 50$[13,17,19,21,24,25]), questionable definitions and classification criteria used for diagnostic groups (eg, etiology of MCI unknown, no confirmation of underlying neurodegeneration[15–17]), combining different clinical groups (eg, cognitively healthy and MCI subjects[20]), potential confounders affecting statistical analyses that were not accounted for,[19] and varied definitions and methods to calculate PE[49]. Due to these overall differences as well as the heterogeneity of study design, participants, and measures, we were not able to conduct a formal meta-analysis or include a funnel plot. However, it should be noted that the potential for publication bias, with respect to statistically significant results for the themes identified in this study, is likely to be high. That is, if we assume that a lack of PE is associated with cognitive decline, and that on average the 27 studies were just adequately powered (ie, had an 80% probability) to detect an association between lack of PE and cognitive decline, we would have only a 7% probability of observing ≥25 of 27 studies returning a significant effect. Moreover, the probability of observing >25 of 27 studies with a positive effect would not reach about 50% until the average power of the 27 studies exceeds 90%. We do not know the power of each of the 27 published studies, but an average of 90% seems implausible. Thus, it would be more plausible that the number of studies performed is higher than 27, with all of the other (non-published) studies not submitted because they had "negative" findings. In this instance, the published literature may overrepresent statistically significant findings.

It should also be considered that, as our primary focus was on PEs as a potential marker for cognitive decline, we included only articles that were assigned to Category C. Therefore, we excluded several articles that focused on understanding mechanisms underlying PEs, for example, by determining whether PEs can be attributed to repeated content rather than context effects,[50] or by investigating different aspects of memory (eg, encoding vs. retrieval) as underlying mechanisms of PEs.[51] We note that those and other articles were not overlooked in the review process, but rather fell outside the scope of the current review. On the other hand, novel, potentially eligible, Category C articles may have been published while the current article was in preparation (eg[52,53]), but were not included as they were not available online before our final search date (January 2020). Although these two articles were not included in this review, they provide further support for the hypothesis that the absence of PEs is a potential marker of cognitive decline.

This review provides an important stepping stone for future work on PE as a useful marker in AD research. To our knowledge, this is the first attempt to summarize the PE literature with a focus on PEs as a potential indicator of subsequent risk rather than as a source of bias in estimating group-level mean cognitive trajectories or as a confounding variable in determining whether an individual patient's cognitive function has changed on a subsequent testing occasion. We performed a comprehensive, systematic literature search followed by a thorough review process benefiting from a multidisciplinary team with expertise in cognitive aging, psychometrics, neuropsychology, medicine, and library science. This study thus offers a novel valuable perspective on the concept and implications of PEs in the study of cognitive aging and dementia. Most of the included articles supported the idea that lower PEs could reflect subtle learning deficits and thereby represent an early clinical symptom of AD.

The idea that an absence of PEs could indicate (subtle) cognitive impairment, implies that researchers should consider characterizing PEs as a marker or risk factor of cognitive decline in studies of cognitive aging and AD. For example, identifying individuals with lower than expected PEs could potentially serve as a cost-effective strategy to enrich enrollment in longitudinal studies, and predict who might be at higher risk for developing AD. Furthermore, identifying high-risk individuals may be an effective strategy to enroll high-risk individuals in randomized-controlled trials of disease-modifying therapeutics. To develop evidence-based recommendations for enrollment in clinical trials, additional research is needed to determine the possible role of lower PEs for estimating AD risk, and potentially in combination with other approaches such as AD biomarkers. We did not find any evaluations of PEs with respect to longitudinal imaging or fluid biomarker data, so it will be important to relate the magnitude of PEs to longitudinal neuroimaging and biomarker data in future research. Furthermore, methods for calculating PEs also warrant consideration in future work, as we found that those methods varied widely, which may have accounted, in part, for contrasting findings across studies. Using more sophisticated characterization of PEs across specific cognitive domains and applying modern psychometric techniques to develop reliable estimates of learning and practice could improve identification of early stage AD. This, in turn, will help to clarify whether individuals with lower PEs are at higher risk for conversion to AD.

In conclusion, we found accumulating evidence that a lack of PEs may represent an early indicator of future cognitive decline and that lower PEs are associated with specific AD biomarkers. The combination of PEs and these biomarkers may yield an optimal approach to estimate AD risk. Future research could then focus on identifying high-risk cohorts from a combination of cognitive and clinical features, and lower than expected PEs. This cost-effective strategy may be able to enrich samples in clinical trials and provide a valuable cognitive marker for subtle pharmacological or treatment responses.

ACKNOWLEDGMENTS

This work was supported by National Institutes of Health (NIH) grant R01AG029672 (Paul K. Crane, PI) and by funding from NIH-National Institute of General Medicine SC3.GM122662 (Nancy S. Foldi, PI).

CONFLICT OF INTEREST

The authors declare that they have no conflict of interest.

REFERENCES

1. Salthouse TA. Robust cognitive change. *J Int Neuropsychol Soc*. 2012;18(4):749-756
2. Rabbitt P, Diggle P, Smith D, Holland F, Mc Innes L. Identifying and separating the effects of practice and of cognitive ageing during a large longitudinal study of elderly community residents. *Neuropsychologia*. 2001;39(5):532-543
3. Duff K. Evidence-based indicators of neuropsychological change in the individual patient: relevant concepts and methods. *Arch Clin Neuropsychol*. 2012;27(3):248-261
4. McKhann GM, Knopman DS, Chertkow H, et al. The diagnosis of dementia due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers Dement*. 2011;7(3):263-269
5. American Psychiatric Association, Diagnostic and statistical manual of mental disorders (DSM-5®). 2013: American Psychiatric Pub
6. Goldberg TE, Harvey PD, Wesnes KA, Snyder PJ, Schneider LS. Practice effects due to serial cognitive assessment: implications for preclinical Alzheimer's disease randomized controlled trials. *Alzheimers Dement (Amst)*. 2015;1(1):103-111
7. Beglinger LJ, Gaydos B, Tangphao-Daniels O, et al. Practice effects and the use of alternate forms in serial neuropsychological testing. *Arch Clin Neuropsychol*. 2005;20(4):517-529
8. Duff K, Lyketsos CG, Beglinger LJ, et al. Practice effects predict cognitive outcome in amnestic mild cognitive impairment. *Am J Geriatr Psychiatry*. 2011;19(11):932-939
9. Hassenstab J, Ruvolo D, Jasielec M, Xiong C, Grant E, Morris JC. Absence of practice effects in preclinical Alzheimer's disease. *Neuropsychology*. 2015;29(6):940-948
10. Jessen F, Amariglio RE, van Boxtel M, et al. A conceptual framework for research on subjective cognitive decline in preclinical Alzheimer's disease. *Alzheimers Dement*. 2014;10(6):844-852
11. Rentz DM, Parra Rodriguez MA, Amariglio R, Stern Y, Sperling R, Ferris S. Promising developments in neuropsychological approaches for the detection of preclinical Alzheimer's disease: a selective review. *Alzheimers Res Ther*. 2013;5(6):58
12. Calamia M, Markon K, Tranel D. Scoring higher the second time around: meta-analyses of practice effects in neuropsychological assessment. *Clin Neuropsychol*. 2012;26(4):543-570
13. Duff K, Horn KP, Foster NL, Hoffman JM. Short-term practice effects and brain hypometabolism: preliminary data from an FDG PET study. *Arch Clin Neuropsychol*. 2015;30(3):264-270
14. Oltra-Cucarella J, Sánchez-SanSegundo M, Ferrer-Cascales R, Initiative Alzheimer's disease neuroimaging. Cognition or genetics? Predicting Alzheimer's disease with practice effects, APOE genotype, and brain metabolism. *Neurobiol Aging*. 2018;71:234-240
15. Cooper DB, Epker M, Lacritz L, et al. Effects of practice on category fluency in Alzheimer's disease. *Clin Neuropsychol*. 2001;15(1):125-128
16. Cooper DB, Lacritz LH, Weiner MF, Rosenberg RN, Cullum CM. Category fluency in mild cognitive impairment: reduced effect of practice in test-retest conditions. *Alzheimer Dis Assoc Disord*. 2004;18(3):120-122
17. Darby D, Maruff P, Collie A, McStephen M. Mild cognitive impairment can be detected by multiple assessments in a single day. *Neurology*. 2002;59(7):1042-1046
18. Dodge HH, Wang C-N, Chang C-CH, Ganguli M. Terminal decline and practice effects in older adults without dementia: the MoVIES project. *Neurology*. 2011;77(8):722-730
19. Duff K, Anderson JS, Mallik AK, et al. Short-term repeat cognitive testing and its relationship to hippocampal volumes in older adults. *J Clin Neurosci*. 2018;57:121-125
20. Duff K, Beglinger LJ, Moser DJ, Paulsen JS, Schultz SK, Arndt S. Predicting cognitive change in older adults: the relative contribution of practice effects. *Arch Clin Neuropsychol*. 2010;25(2):81-88
21. Duff K, Beglinger LJ, Schultz SK, et al. Practice effects in the prediction of long-term cognitive outcome in three patient samples: a novel prognostic index. *Arch Clin Neuropsychol*. 2007;22(1):15-24
22. Duff K, Beglinger LJ, Van Der Heiden S, et al. Short-term practice effects in amnestic mild cognitive impairment: implications for diagnosis and treatment. *Int Psychogeriatr*. 2008;20(5):986-999
23. Duff K, Chelune G, Dennett K. Within-session practice effects in patients referred for suspected dementia. *Dement Geriatr Cogn Disord*. 2012;33(4):245-249
24. Duff K, Foster NL, Hoffman JM. Practice effects and amyloid deposition: preliminary data on a method for enriching samples in clinical trials. *Alzheimer Dis Assoc Disord*. 2014;28(3):247
25. Duff K, Hammers DB, Dalley BCA, et al. Short-term practice effects and amyloid deposition: providing information above and beyond baseline cognition. *J Prev Alzheimers Dis*. 2017;4(2):87-92
26. Galvin JE, Powlishta KK, Wilkins K, et al. Predictors of preclinical Alzheimer disease and dementia: a clinicopathologic study. *Arch Neurol*. 2005;62(5):758-765
27. Hanyu H, Sakurai H, Hirao K, Sato T, Iwamoto T. Difference in practice effects depending on cerebral perfusion pattern in mild cognitive impairment. *Int J Geriatr Psychiatry*. 2008;23(1):111-112
28. Vivot A, Power MC, Maria Glymour M, et al. Jump, hop, or skip: modeling practice effects in studies of determinants of cognitive change in older adults. *Am J Epidemiol*. 2016;183(4):302-314
29. Ihara R, Iwata A, Suzuki K, et al. Clinical and cognitive characteristics of preclinical Alzheimer's disease in the Japanese Alzheimer's disease neuroimaging initiative cohort. *Alzheimers Dement (N Y)*. 2018;4:645-651
30. Jonaitis EM, Koscik RL, La Rue A, Johnson SC, Hermann BP, Sager MA. Aging, practice effects, and genetic risk in the Wisconsin registry for Alzheimer's prevention. *Clin Neuropsychol*. 2015;29(4):426-441
31. Machulda MM, Hagen CE, Wiste HJ, et al. Practice effects and longitudinal cognitive change in clinically normal older adults differ by Alzheimer imaging biomarker status. *Clin Neuropsychol*. 2017;31(1):99-117
32. Machulda MM, Shane Pankratz V, Christianson TJ, et al. Practice effects and longitudinal cognitive change in normal aging vs. incident mild cognitive impairment and dementia in the Mayo clinic study of aging. *Clin Neuropsychol*. 2013;27(8):1247-1264
33. Sánchez-Benavides G, Gispert JD, Fauria K, Molinuevo JL, Gramunt N. Modeling practice effects in healthy middle-aged participants of

the Alzheimer and families parent cohort. *Alzheimers Dement (Amst)*. 2016;4:149-158
34. Schrijnemaekers AMC, de Jager CA, Hogervorst E, Budge MM. Cases with mild cognitive impairment and Alzheimer's disease fail to benefit from repeated exposure to episodic memory tests as compared with controls. *J Clin Exp Neuropsychol*. 2006;28(3):438-455
35. Suchy Y, Kraybill ML, Franchow E. Practice effect and beyond: reaction to novelty as an independent predictor of cognitive decline among older adults. *J Int Neuropsychol Soc*. 2011;17(1):101-111
36. Wilson RS, Capuano AW, Yu L, et al. Neurodegenerative disease and cognitive retest learning. *Neurobiol Aging*. 2018;66:122-130
37. Zehnder AE, Bläsi S, Berres M, Spiegel R, Monsch AU. Lack of practice effects on neuropsychological tests as early cognitive markers of Alzheimer disease?. *Am J Alzheimers Dis Other Demen*. 2007;22(5):416-426
38. Gross AL, Inouye SK, Rebok GW, et al. Parallel but not equivalent: challenges and solutions for repeated assessment of cognition over time. *J Clin Exp Neuropsychol*. 2012;34(7):758-772
39. Bläsi S, Zehnder AE, Berres M, Taylor KI, Spiegel R, Monsch AU. Norms for change in episodic memory as a prerequisite for the diagnosis of mild cognitive impairment (MCI). *Neuropsychology*. 2009;23(2):189
40. Cacciamani F, Salvadori N, Eusebi P, et al. Esvidence of practice effect in CANTAB spatial working memory test in a cohort of patients with mild cognitive impairment. *Applied Neuropsychology: Adult*. 2018;25(3):237-248
41. Duff K. One-week practice effects in older adults: tools for assessing cognitive change. *Clin Neuropsychol*. 2014;28(5):714-725
42. Gavett BE, Ashendorf L, Gurnani AS. Reliable change on neuropsychological tests in the Uniform Data Set. *J Int Neuropsychol Soc*. 2015;21(7):558-567
43. Chelune GJ, Duff K. The Assessment of Change: Serial Assessments in Dementia Evaluations. In: Ravdin LD.Katzen HL, eds. *Handbook on the Neuropsychology of Aging and Dementia*. Cham, Switzerland: Springer; 2019:61-76
44. Heaton RK, Temkin N, Dikmen S, et al. Detecting change: a comparison of three neuropsychological methods, using normal and clinical samples. *Arch Clin Neuropsychol*. 2001;16(1):75-91
45. Van der Elst W, Van Boxtel MPJ, Van Breukelen GJP, Jolles J. Detecting the significance of changes in performance on the Stroop Color-Word Test, Rey's Verbal Learning Test, and the Letter Digit Substitution Test: the regression-based change approach. *J Int Neuropsychol Soc*. 2008;14(1):71-80
46. Goldberg TE, Keefe RSE, Goldman RS, Robinson DG, Harvey PD. Circumstances under which practice does not make perfect: a review of the practice effect literature in schizophrenia and its relevance to clinical treatment studies. *Neuropsychopharmacology*. 2010;35(5):1053-1062
47. Jack CR, Jr, Bennett DA, Blennow K, et al. NIA-AA research framework: toward a biological definition of Alzheimer's disease. *Alzheimers Dement*. 2018;14(4):535-562
48. Racine AM, Gou Y, Fong TG, et al. Correction for retest effects across repeated measures of cognitive functioning: a longitudinal cohort study of postoperative delirium. *BMC Med Res Method*. 2018;18(1):69
49. Howieson DB, Carlson NE, Milar Moore M, et al. Trajectory of mild cognitive impairment onset. *J Int Neuropsychol Soc*. 2008;14(2):192-198
50. Gross AL, Chu N, Anderson L, Maria Glymour M, Jones RN, Coalition Against Major Diseases. Do people with Alzheimer's disease improve with repeated testing? Unpacking the role of content and context in retest effects. *Age Ageing*. 2018;47(6):866-871
51. Stamate A, Logie RH, Baddeley AD, Della Sala S. Forgetting in Alzheimer's disease: is it fast? Is it affected by repeated retrieval?. *Neuropsychologia*. 2020;138:107351
52. Wang G, Kennedy RE, Goldberg TE, Fowler ME, Cutter GR, Schneider LS. Using practice effects for targeted trials or sub-group analysis in Alzheimer's disease: how practice effects predict change over time. *PLoS One*. 2020;15(2):e0228064
53. Hammers DB, Suhrie KR, Dixon A, Porter S, Duff K. Reliable change in cognition over 1 week in community-dwelling older adults: a validation and extension study. *Arch Clin Neuropsychol*. 2020. doi: https://doi.org/10.1093/arclin/acz076

How to cite this article: Jutten RJ, Grandoit E, Foldi NS, et al. Lower practice effects as a marker of cognitive performance and dementia risk: A literature review. *Alzheimer's Dement*. 2020;12:e12055. https://doi.org/10.1002/dad2.12055