



Pearson

# Efficacy Reporting Framework

14 August 2020



## Introduction


In 2013, we were the first company to make a commitment to measure our impact on some of the outcomes that matter most to learners, such as academic achievement. But there was no rule book and no model to follow. We've had to carve our own path to define what efficacy looks like in education.

While our approach is rigorous, the concept underlying it is simple: we use evidence and research to design products and solutions to help learners achieve the outcomes that matter to them. Then, we measure the impact of using our products, report that impact in a transparent way, and use what we learn to help learners – and ourselves – continuously improve.

Today, we are taking what we have learned and evolving our approach. We are focusing more on designing products to have a measurable impact, not just during education, but on employability and lifelong learning as well. We want our commitment to efficacy to be a reason for learners to believe in Pearson, to see us as their trusted guide to lifelong learning, as they navigate a changing world of work. And now that the idea of a job for life is gone, people need to continuously grow, demonstrate their skills and adapt their talent.

People need a lifetime of learning and so we must refocus and redesign learning. The way we learn needs to support the development of the key skills people need to thrive today and in the future.

Our efficacy reports help us, and the wider education community, build a better understanding of not just what works, but how, why, and in what context — helping us learn, not guess, about how the design and use of products relates to the achievement of outcomes that matter most.



### **Kate Edwards**

Senior Vice President,  
Efficacy and Learning Research, Pearson  
August, 2020

### **Special thanks**

We want to thank all the customers, test-takers, research institutions and organizations we have collaborated with to date. If you are interested in partnering with us on future efficacy research, have feedback or suggestions for how we can improve, or want to discuss your approach to using or researching our products, we would love to hear from you at [efficacy@pearson.com](mailto:efficacy@pearson.com).

# Contents

- 04 About this document
- 05 Part 1: Overview of the Efficacy Reporting Framework
  - 05 Background to efficacy reporting
  - 09 The efficacy reporting process
  - 10 Creating accurate efficacy statements
- 11 Part 2: The Efficacy Reporting Framework in detail
  - 11 Process 1: Learner outcomes
  - 12 Process 2: Study preparation
  - 14 Process 3: Study write-up and draft efficacy statements
  - 16 Process 4: Product Efficacy Reports
  - 19 Efficacy statement terminology by research design type
  - 26 Audit of efficacy statements

## About this document

This document explains how Pearson approaches efficacy reporting.

**Part 1** explains what we hope to achieve by our efficacy reporting and how we have designed our process to achieve this. Anyone interested in evaluating our process and the rigor behind our findings should find this section useful.

**Part 2** goes into detail about the activities we perform and the documentation we use at every stage of the efficacy reporting process. This section also includes the complete guidance we follow when preparing evidence-based efficacy statements, and a breakdown of the work carried out by our auditors. Anyone who is thinking of following a similar process to prepare their own efficacy reporting, or who is simply in search of more granular detail than we provide in Part 1, should find this section useful.

We refer to our approach to efficacy reporting — including the process, controls, documentation (such as our guidance on efficacy statement terminology by research design type), reviews and third-party audit — as the Pearson Efficacy Reporting Framework.

We are making the details of the Efficacy Reporting Framework public in the interests of transparency, and in the hope that sharing our approach will encourage others to give feedback about how we can improve it, or replicate and build on what we have done. This document is published under a Creative Commons licence, so you may share and use it freely as long as you reference it in any work you produce as a result. If you do decide to follow our process, and you want to discuss any aspect of it with us, please contact [efficacy@pearson.com](mailto:efficacy@pearson.com).

# Part 1: Overview of the Efficacy Reporting Framework

## Background to efficacy reporting

### Why we publish efficacy reporting

At Pearson, we recognize how important it is to understand the relationships between the design and use of our products and the outcomes that matter most to students and educators. We want to understand, not only what works, but also for whom and why.

We engage in research on our products to improve our understanding in these areas. Our focus on improving the efficacy of our products, by applying evidence-based insight both from learning research and design and from impact evaluation research, allows us to relentlessly improve our products so they can support delivery of better outcomes for more learners.

In 2013, we made a public commitment to begin regularly publishing the results of this research by 2018. Rigorous impact evaluation research and transparent reporting allow us to help more learners, learn more by:

- better understanding and highlighting relationships between the use of our products and the learning and learner outcomes that matter most to students and educators, and using this understanding to help others achieve comparable or even better impact themselves
- sharing the evidence underpinning the design and development of new and existing products
- galvanizing other learning companies to follow suit and measure their impact by the outcomes they deliver for learners

## Our journey

Year	Action
2013	<p>Pearson commits to start publishing annual reports about the efficacy of its products, and to have those reports externally audited by 2018.</p> <p>Pearson publishes <b><u>The incomplete guide to delivering learning outcomes</u></b> with the Shared Value Initiative.</p> <p>Pearson publishes <b><u>From good intentions to real impact</u></b>, an initial approach to efficacy research and reporting developed in partnership with Nesta.</p>
2014	<p>Pearson appoints PricewaterhouseCoopers LLP (PwC) as its independent auditor.</p>
2014–2016	<p>Pearson refines its approach with input from organizations like the American Educational Research Association, What Works Clearinghouse, and the Efficacy Academic Network.</p>
2015	<p>The Efficacy Academic Network is formed. It is made up of four leading academics from the US, UK, and Australia: Eva Baker, Philippa Cordingley, Chris Dede, and Gordon Stanley.</p> <p>The first dry run of the efficacy reporting process: Pearson publishes reports on five of its products.</p>
2016	<p>The second dry run of the efficacy reporting process: Pearson goes deeper and broader with efficacy reporting — looking at more products and detailing more rigorous research.</p>
2017	<p>The third dry run of the efficacy reporting process: Pearson publishes five reports and this time also subjects them to a mock audit.</p> <p>Pearson examines its approach in more detail with SRI Education, with the results published as <b><u>Understand, implement, evaluate</u></b>.</p> <p>Pearson sponsors and contributes to the EdTech Efficacy Research Academic Symposium, fostering and contributing to discussion and debate around efficacy.</p>
2018	<p>Pearson publishes its Efficacy Reporting Framework and transparent, rigorous reports about the efficacy of its most widely used digital products. The Product Efficacy Reports were independently assured by PwC.</p>
2019	<p>Pearson uses feedback from 2018 to improve its approach to efficacy reporting — streamlining the process, formalizing and strengthening the role of independent peer reviewers, and publishing Educator Guides alongside the efficacy reports, to provide additional insight and support for implementation.</p> <p>Building on its 2013 commitment, Pearson publishes three new independently assured reports about the efficacy of its products.</p>
2020	<p>Pearson publishes various reports (in addition to independently assured reports) focusing on products designed to support individuals to be successful in their careers.</p>

## What efficacy reporting looks like

Key findings from Pearson's research are reported in the form of efficacy statements. These are evidence-based statements summarizing the primary conclusions of a study. Each statement relates to the use of a Pearson product and/or its impact on learners and learning.

Each efficacy statement represents a finding from a specific study using a particular type of research design. The effects described in an efficacy statement are dependent on the research design used and its limitations. For that reason, they should be interpreted within the context of the research design used that supports the evidence and details how, when, and where the evidence was collected.

To be as open and transparent as possible about how we design, develop, and evaluate the impact of use of our products on learning, Pearson publishes two main kinds of efficacy reports. Each kind of report provides a different level of insight into the research supporting the efficacy statements relating to the use of the product.

**Technical Research Reports** each describe a single piece of research into the use of a product, undertaken or commissioned by Pearson to meet the standards expected for publication in peer-reviewed academic journals. Technical Research Reports include enough detail for another researcher to replicate the research.

**Product Efficacy Reports** summarize all the relevant research related to the use of a single product, based on one or more Pearson Technical Research Reports, as well as any external studies we discover that meet our criteria for inclusion. They also include information about the learning research that informed the product's design, and how the product is designed to be used.

## Types of research design for efficacy reporting

The research that underpins our audited efficacy statements includes correlational designs, comparative designs, and causal designs.

- **Correlational designs** study students using Pearson products to investigate whether there is a relationship between how they use the product and its intended outcomes for learners, while controlling for background factors like baseline achievement.
- **Comparative designs** compare learner outcomes for students who use a Pearson product to those of students who don't, but those student groups may be very different from one another, which means we cannot rule out confounding factors.
- **Causal designs** compare learner outcomes for students who use a Pearson product to those of students who don't, with student groups that are highly similar at baseline, allowing us to rule out as many confounding factors as possible.

It should be noted that Pearson makes five types of efficacy statements derived from these three types of research design, as explained on page 19. In addition to correlational, comparative with caveats, comparative, and causal efficacy statements, we also make descriptive efficacy statements, which can be derived from any study capable of supporting another type of efficacy statement. Descriptive efficacy statements are purely statements about the sample, not any wider population.

Pearson makes use of a wider range of research activity types than this throughout the lifecycle of a product, including market research and exploratory learning research. But when we prepare efficacy reporting and audited efficacy statements for a product, we use these three specific types of study design.

This selection of study design types for our research allows us to build a body of evidence over time, and to use the evidence to demonstrate impact to our customers, improve existing products, and develop new ones.

## Reviewing and auditing efficacy reports

It is vitally important that our customers know they can trust the statements we make about the use of our products.

A team of researchers from SRI Education (a third-party validator) with expertise in study design and quantitative methodology independently reviews Pearson's research and efficacy statements relating to the use of products and their impact on learners and learning.

First, the team helped Pearson codify its research quality criteria into tiered levels of evidence to reflect best practices in the field. Criteria for these tiered levels of evidence drew both from the tiered evidence levels in the USA's Every Student Succeeds Act (ESSA) and the 2006 American Educational Research Association's Standards for Reporting on Empirical Social Science Research.

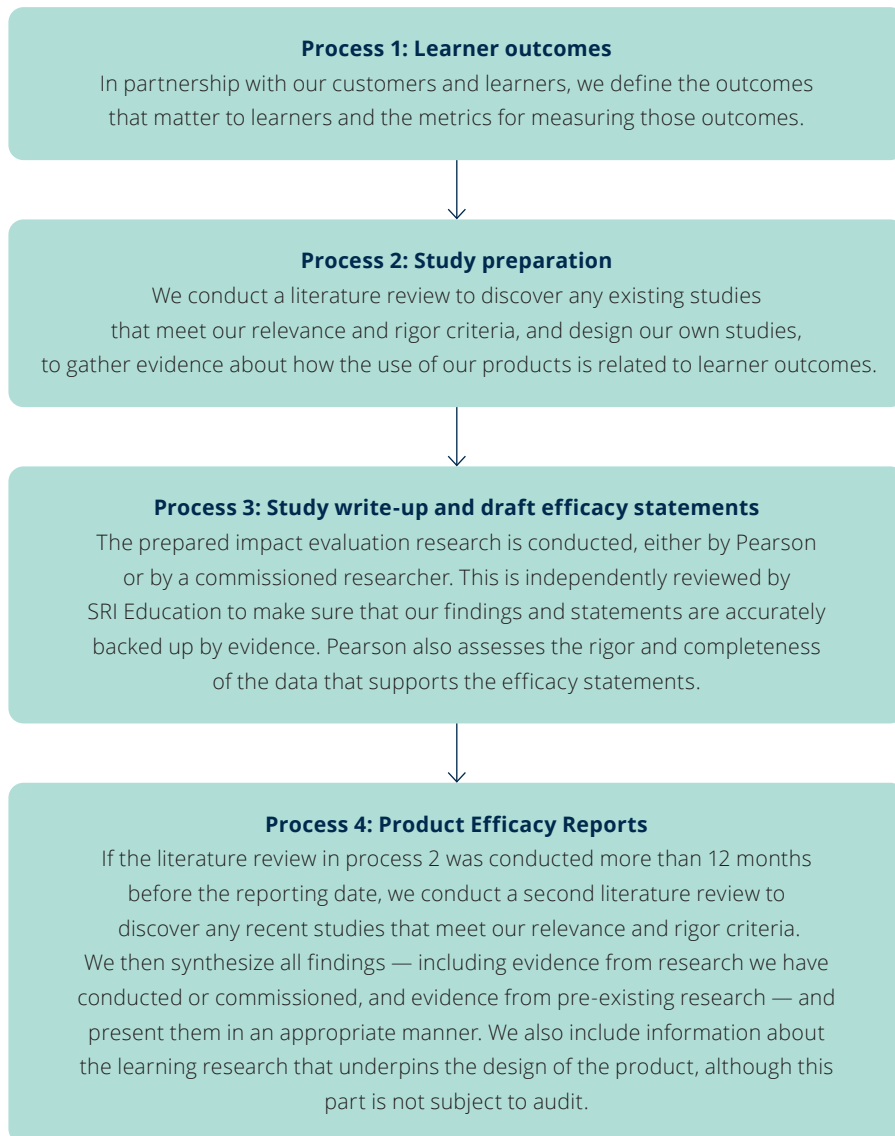
Then the team assesses whether Pearson's analyses, technical reporting, and associated efficacy statements complied with these research quality criteria.

We have appointed PricewaterhouseCoopers LLP (PwC) to provide limited assurance on the efficacy statements set out in our Product Efficacy Reports. Each Product Efficacy Report includes an independent assurance report from PwC on the efficacy statements about the product. This is to demonstrate that the statements accurately reflect the research that has been carried out. During their work, as well as the efficacy statements themselves, PwC considers the design and operation of the process we use to generate the statements.



# The efficacy reporting process

Pearson has developed a four-stage process to prepare reporting on the impact of use of our products on outcomes.



The process includes controls designed to make sure our reporting of the research we conduct is rigorous and accurate. Each stage of the process also has a set of associated review documents, or checklists. These documents:

- ensure that key activities are applied consistently across all products
- form a record of the key activities completed
- inform improvements to the product in the future

The prominent academics of the Efficacy Academic Network provided feedback during the development of the efficacy reporting process. This feedback, the independent assurance from PwC, and the independent review by SRI Education are all ways in which we improve the rigor of our efficacy reporting.

## Creating accurate efficacy statements

We need our customers to know they can trust what we tell them about our products. It is vitally important, therefore, that our efficacy statements accurately represent the evidence from the research they are based on.

This means the wording of efficacy statements must be carefully considered to make sure it is aligned with the standard and design of the research it is derived from. We make five types of efficacy statements. Three of these map directly to the three research design types.

- **Descriptive** efficacy statements can correspond to correlational, comparative, or causal designs.
- **Correlational** efficacy statements correspond to correlational designs.
- **Comparative** efficacy statements correspond to comparative designs.
- **Comparative with caveats** efficacy statements correspond to comparative designs.
- **Causal** efficacy statements correspond to causal designs.

For more detail on the different types of research design, specific research standards applicable to each research design, and example efficacy statements, please see page 19.

## Part 2: The Efficacy Reporting Framework in detail

### Process 1: Learner outcomes

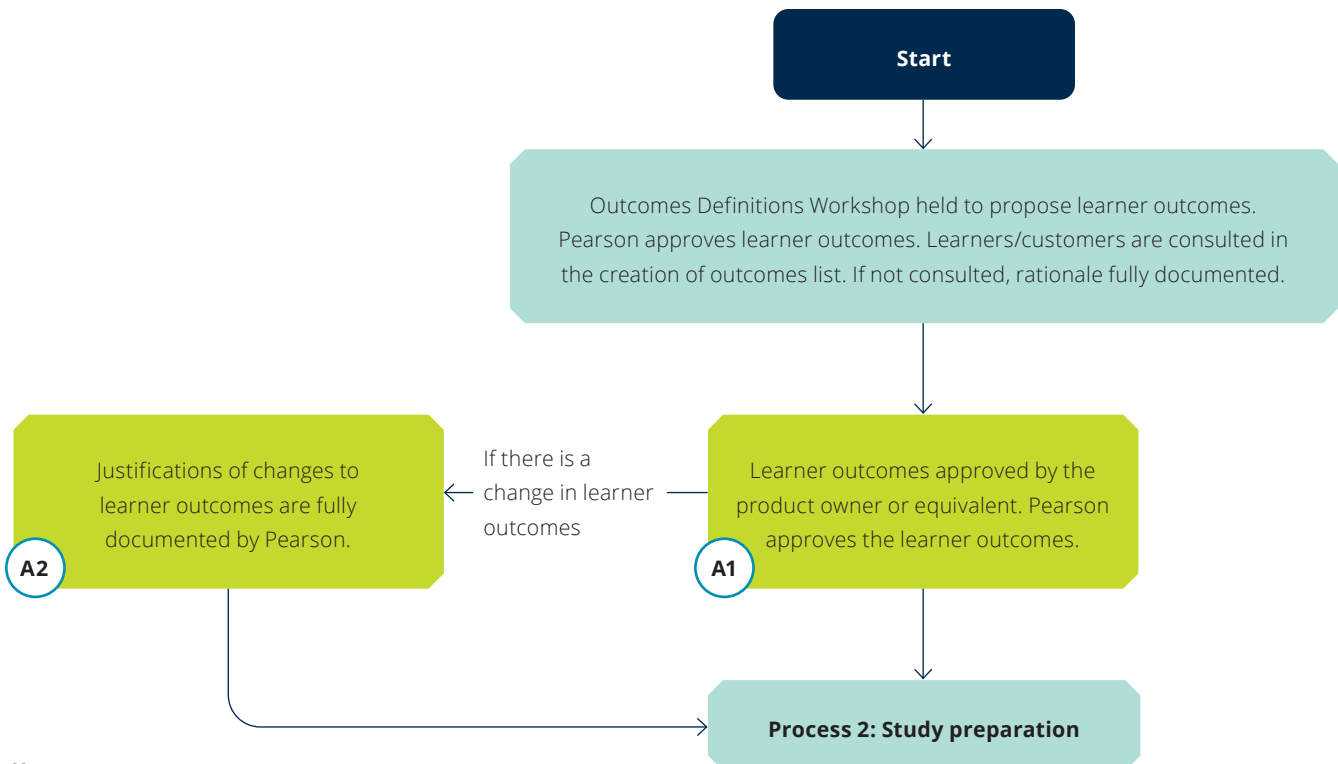
In process 1, in partnership with customers and learners, Pearson defines the outcomes that matter to learners and the metrics for measuring those outcomes.

#### Artefacts required in process 1

Reference	Artefact	Purpose
A1	Learner outcomes form, including evidence of review and approval by Pearson	Documents the agreed intended learner outcomes for a product, as well as the metrics and leading indicators that can be used to measure them
A2	Learner outcomes definition sign-off and change log, documenting rationales for all changes to learner outcomes, with supporting documents and evidence of review and approval by Pearson	Documents changes made to the product's learner outcomes over time

#### What's involved

##### Process 1: Learner outcomes



#### Key

- Relevant process for audit
- Relevant control for audit
- A Relevant artefact for audit

## Process 2: Study preparation

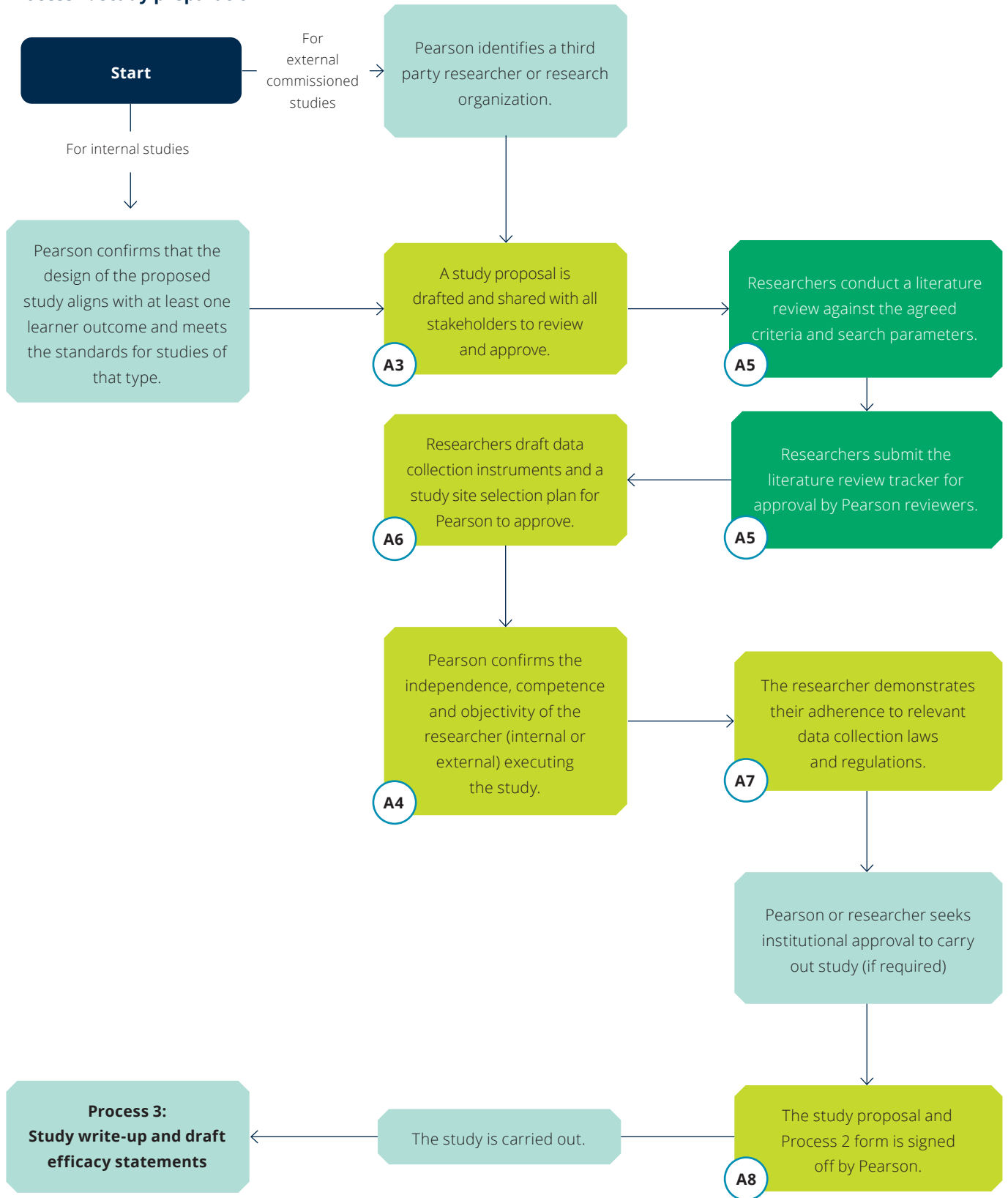
In process 2, we conduct a literature review to discover any existing studies that meet our relevance and rigor criteria, and design our own studies, to gather evidence about how the use of our products is related to learner outcomes.

### Artefacts required in process 2

Reference	Artefact	Purpose
A3	Study proposal	Details a proposed impact evaluation research study
A4	Independence, competence, and objectivity confirmation form for the researcher conducting the research	Provides evidence that the lead researcher is able to conduct the research study objectively
A5	Literature review tracker, with approach, search parameters, and results, by Pearson	Provides evidence of the literature search methods and results in terms of the relevance, and adherence to research quality standards, of the impact evaluation research studies found
A6	Draft data collection instruments	Document any data collection instruments proposed for use in the study
A7	Evidence of compliance before data collection (for example, approved Institutional Review Board application documents)	Provides evidence that the researcher has complied with any data collection rules, or sought relevant data use permissions, before the study is conducted
A8	Signed off process 2 form	Documents sign-off from Pearson team members that all relevant evidence in process 2 has been collected and correctly documented

**What's involved**

**Process 2: Study preparation**



- Key**
- Relevant process for audit
  - Relevant control for audit
  - Control not relevant for audit
  - Relevant artefact for audit

## Process 3: Study write-up and draft efficacy statements

In process 3, the prepared impact evaluation research is conducted, either by Pearson or by a commissioned researcher. This is independently reviewed by a third-party validator (SRI Education) to make sure that our findings and statements are accurately backed up by evidence. Pearson also assesses the rigor and completeness of the data that supports the efficacy statements.

### Artefacts required at process 3

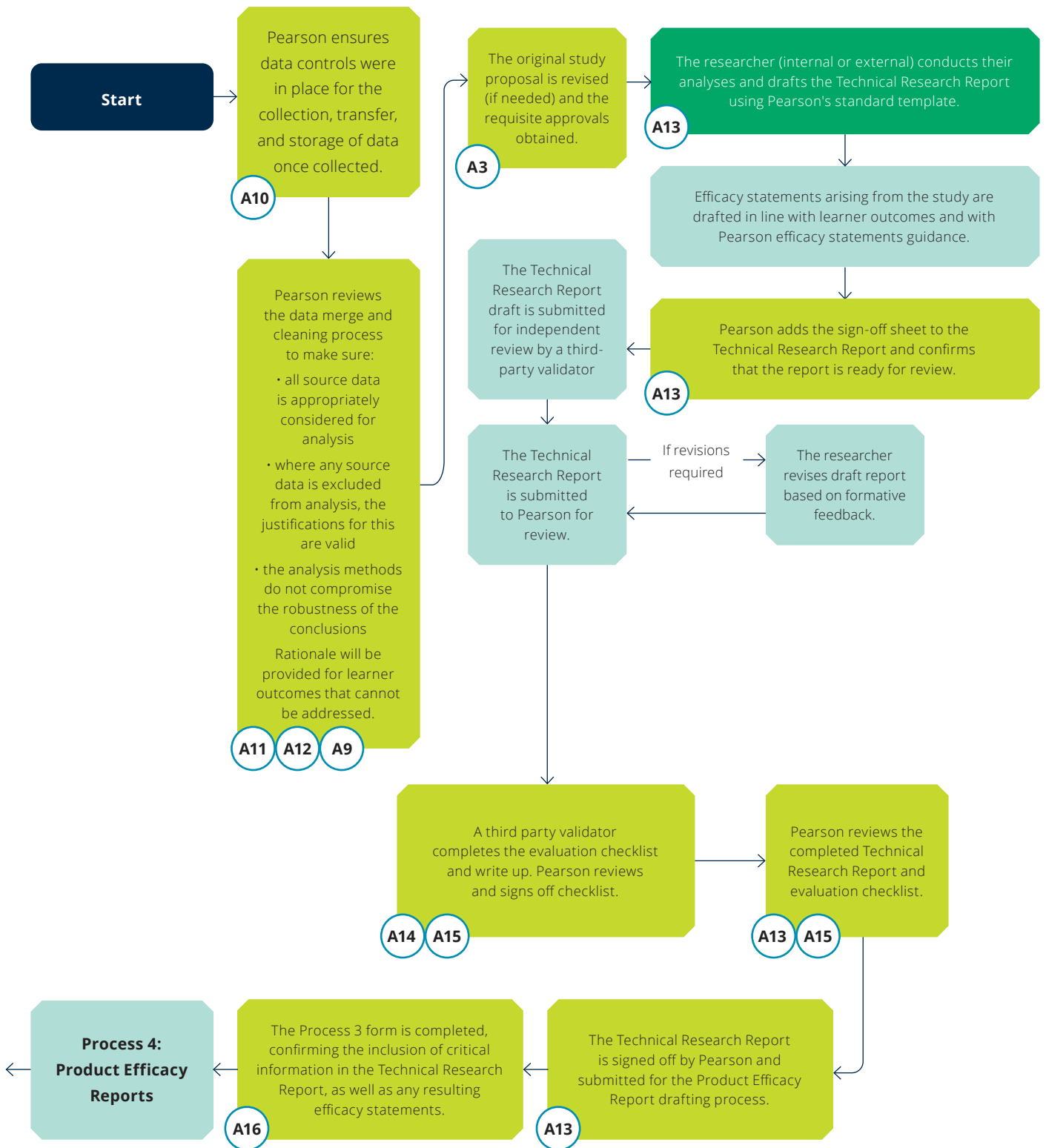
Reference	Artefact	Purpose
A9	Statistical output used to populate the results tables	Documents any statistical codes and outputs used to generate statistical results and tables
A10	Evidence of compliance following data collection (for example, Internal Audit conclusion/report)	Provides evidence to confirm that the researcher complied with any data collection rules stipulated in the Data Use Agreement or Institutional Review Board/ethics board documentation during the process of conducting the study, and that appropriate controls were in place for the collection, transfer, and storage of data
A11	Data merge and clean form, confirming sample sizes after data merging/cleaning and accuracy of the statistical output	Provides evidence that a senior member of Pearson research staff has checked the way the researcher merged and cleaned the data to ensure accuracy and fidelity
A12	Rationale for Evidence Gap form	Provides rationale for learner outcomes that were not addressed
A13	Technical Research Report, including details of the data cleaning process and a sign-off sheet confirming review	Documents the study, including methodology, results, efficacy statements, and any limitations, in enough detail for another researcher to replicate the study
A14	Write-up from independent reviewer, detailing evaluation and confirmation of statistical elements of report, including responses from the researcher that carried out the study	Presents formative feedback from an independent third party (other than PwC), which the Pearson researcher either acts on by updating the Technical Research Report or responds to with a rationale for why the changes were not made
A15	Technical Research Report review checklist, completed and signed off by Pearson and a third-party validator	Documents sign-off by third-party validator that the Technical Research Report meets Pearson's standards for conducting a study, and that any efficacy statements are supported by the research
A16	Signed off process 3 form, confirming inclusion of critical information in the Technical Research Report and any resulting efficacy statements	Documents sign-off from Pearson team members that all relevant evidence in process 3 has been collected and correctly documented

### Artefact resubmissions

Reference	Artefact	Purpose
A3	Study proposal (updated before writing the Technical Research Report), including updated change log showing changes made	Provides evidence and a rationale for any changes that were made to the original study proposal in the course of conducting the study due to unforeseen events (e.g., attrition, missing data, treatment contamination)

**What's involved**

**Process 3: Study write-up and draft efficacy statements**



- Key**
- Relevant process for audit
  - Relevant control for audit
  - Control not relevant for audit
  - A Relevant artefact for audit

## Process 4: Product Efficacy Reports

In process 4, we synthesize all findings from research we have conducted or commissioned, and evidence from pre-existing research — and present them in an appropriate manner. We also include information about the learning research that underpins the design of the product, although this part is not subject to audit.

If the literature review in process 2 was conducted more than 12 months before the reporting date, we also conduct a second literature review in process 4, to discover any recent studies that meet our relevance and rigor criteria.

### Artefacts required in process 4

Reference	Artefact	Purpose
A17	Literature review write-up, describing any additional results from the literature review conducted in process 2; third-party validator to review the write-up	Documents the methods and results of the literature review, for inclusion in the Product Efficacy Report
A18	Signed-off log of all efficacy statements related to the product , with newly added efficacy statements marked out by date added	Provides evidence that all efficacy statements have been agreed and signed off by the relevant stakeholders Provides a rationale for any changes to be made to previously published efficacy statements
A19	Draft Product Efficacy Report	Provides stakeholders (including PwC) with an opportunity to review and feed back on the content of the Product Efficacy Report that summarizes all relevant impact evaluation research related to the use of a product, as well as any external studies found that meet the criteria for inclusion
A20	Final Product Efficacy Report	Summarizes all the relevant impact evaluation research related to the use of a product, as well as any external studies found that meet the criteria for inclusion
A21	Process 4 form, confirming sign-off	Documents sign-off from senior Pearson team members, including the senior vice president of efficacy and research, that all relevant evidence in process 4 has been collected and correctly documented

### Artefact resubmissions

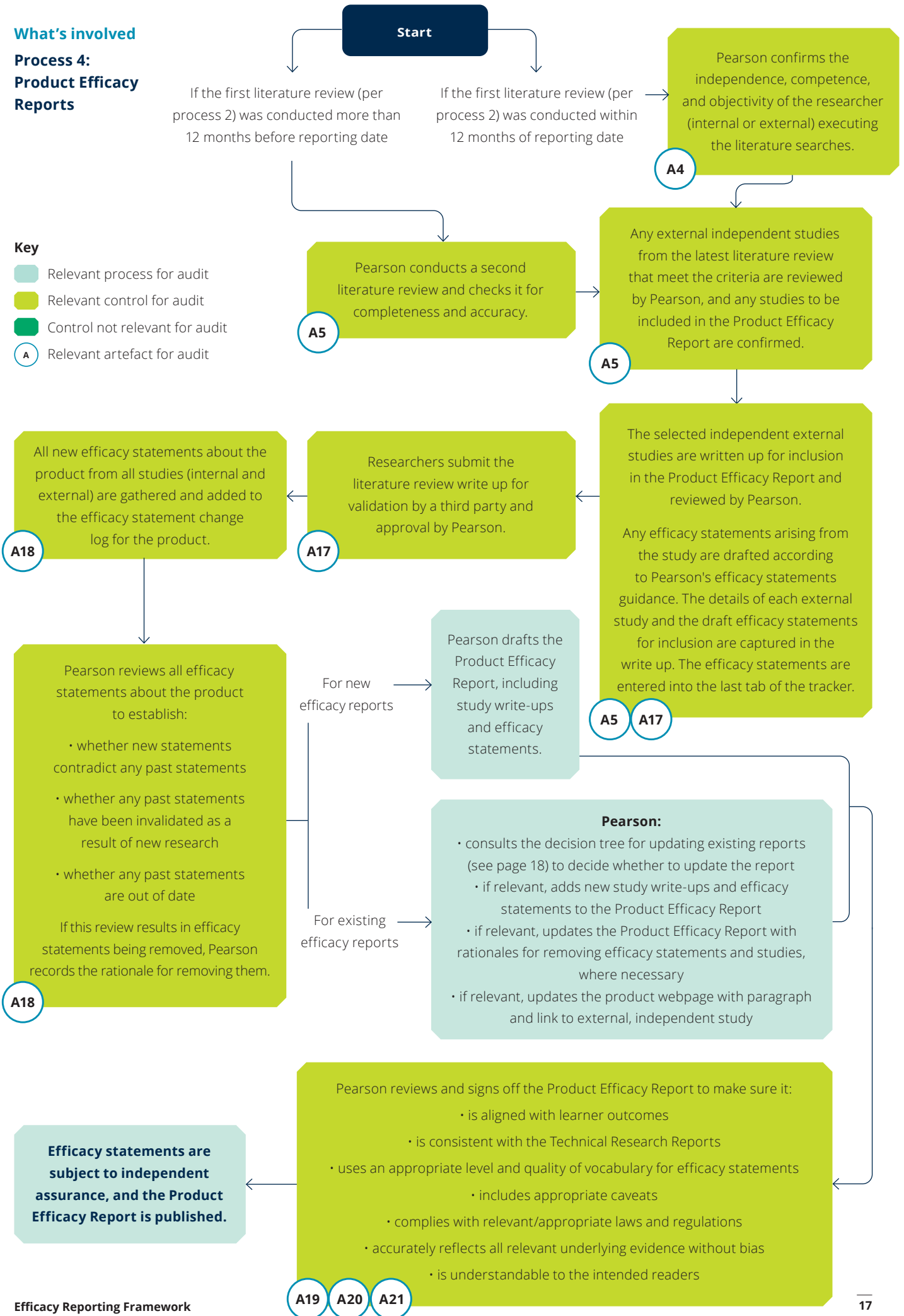
Reference	Resubmitted artefact	Purpose
A5	Literature review tracker, updated with new searches and results, citations for relevant studies, efficacy statements derived from those studies, and confirmation of approval	Documents any additional research that has been published since the last literature review was conducted and is eligible for inclusion in the Product Efficacy Report. If the literature review was conducted more than 12 months before the reporting date, a second literature review is performed, and the results recorded in this tracker



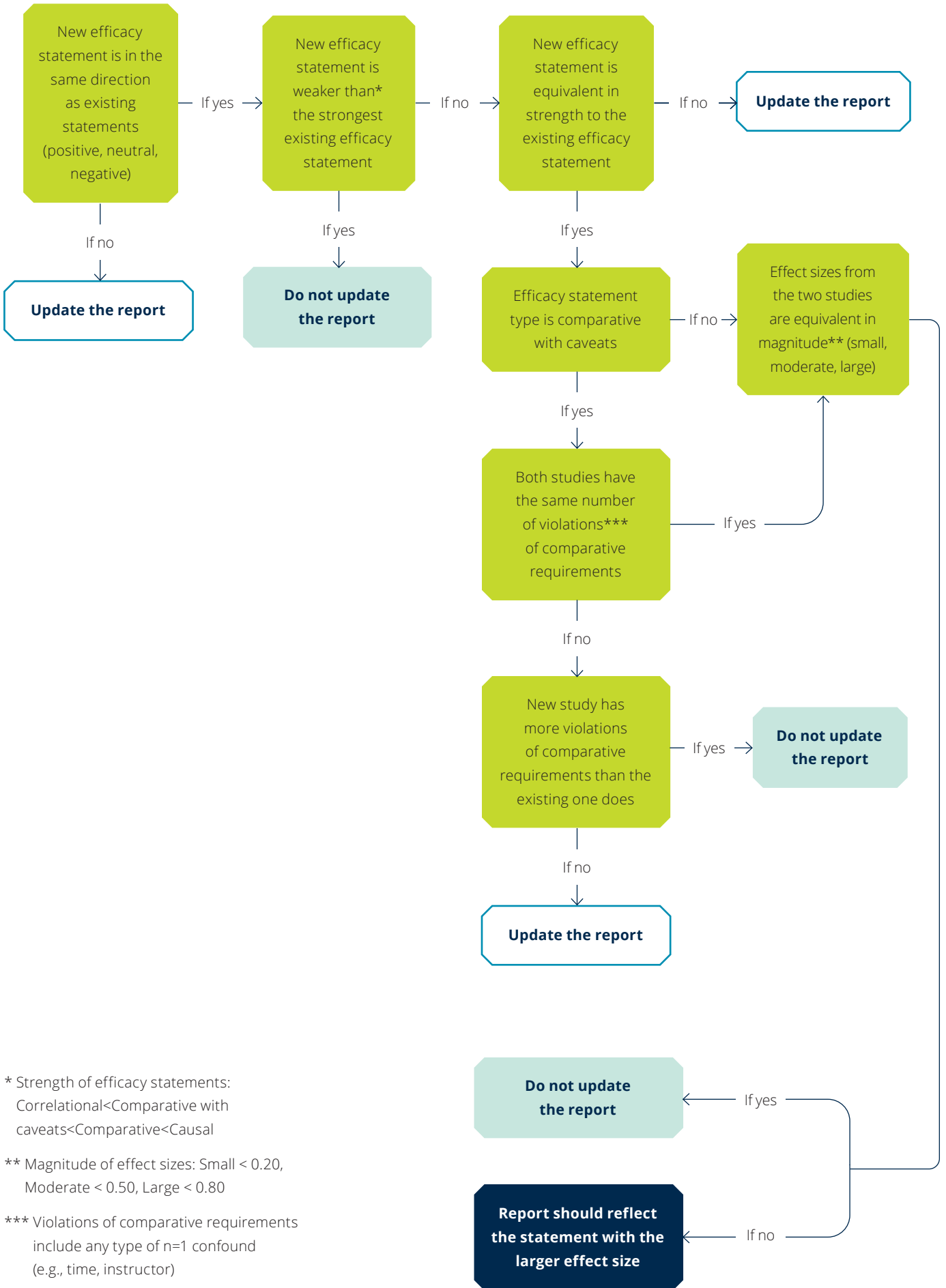
**What's involved**

**Process 4:  
Product Efficacy  
Reports**

- Key**
- Relevant process for audit
  - Relevant control for audit
  - Control not relevant for audit
  - A Relevant artefact for audit



**Decision tree for updating existing reports to reflect new research**



\* Strength of efficacy statements:  
Correlational < Comparative with caveats < Comparative < Causal

\*\* Magnitude of effect sizes: Small < 0.20,  
Moderate < 0.50, Large < 0.80

\*\*\* Violations of comparative requirements include any type of n=1 confound (e.g., time, instructor)

# Efficacy statement terminology by research design type

Pearson uses the following guidance when preparing efficacy statements about the use of its products. The aim of the guidance is to ensure that efficacy statements accurately represent the evidence they are based on, by aligning their wording with the standard and design of the underlying studies.

The standards by which Pearson judges studies are informed by and aligned with recognized research standards and/or frameworks developed by the What Works Clearinghouse and the American Education Research Association.

The illustrative example efficacy statements presented here should not be construed as comprehensive or exhaustive. Rather, they are intended to illustrate the types of wording that would be acceptable for a given combination of efficacy statement type and underlying evidence. The wording of efficacy statements may vary slightly to improve readability in different contexts, as long as our third-party validator agrees that these variations still preserve the original meaning and intent.

Where a study investigated a relationship or difference between factors, Pearson will include efficacy statements that describe the relationship or difference found. Where the study found that the relationship or difference was not statistically significant, the reports will state this clearly.

All efficacy statements should:

- directly address and align with one or more of the research questions
- accurately represent the design and findings aligned with the associated research question(s)
- accurately represent the data, analyses, and findings of the study
- discourage inaccurate interpretations (for example, non-causal efficacy statements should be worded to discourage causal interpretations)

## Types of efficacy statement

Efficacy statements can be categorized as either general or specific. Alongside this broad categorization, efficacy statements may be:

- descriptive
- correlational
- comparative with caveats
- comparative
- causal

In order to make a particular type of efficacy statement, the study must meet the corresponding study design and analysis requirements.

## General versus specific efficacy statements

---

General efficacy statements express a trend, relationship, or effect in simple and broad terms.

Specific efficacy statements provide more detail by:

- communicating more precise numeric estimates
- representing effect sizes or mean differences along with relevant statistical tests
- interpreting general statements in terms of practical significance

Typically, efficacy reporting will use both general and specific efficacy statements.

## Requirements of specific efficacy statements

To be considered complete and accurate, specific efficacy statements must include certain important components.

### Descriptive efficacy statements

---

#### ***Which research designs support descriptive statements?***

Any research design capable of supporting correlational, comparative with caveats, comparative, or causal efficacy statements can also support descriptive statements via analysis of surveys, secondary analysis of administrative data, or cohort analysis with no controls for baseline factors or potential mediating/moderator variables.

#### ***What are the essential requirements of descriptive statements?***

- Result phrased as a specific quantified measure (such as a percentage or student count) supported by the study's analysis
- Refers to the specific study sample to which the statement applies (for example, all respondents to a survey, female students using the product/service, grade levels)
- Refers to a learner outcome of interest
- Language does not express a quantitative correlation or statistical association, or other relationship

#### ***Illustrative examples of descriptive statements***

- Students generally felt the product helped them achieve their goal
- #% of students indicate the product helps them improve their learning “significantly” or “very significantly”

### Correlational efficacy statements

---

#### ***Which research designs support correlational statements?***

Cohort analysis with controls for baseline factors or potential mediating/moderator variables.

#### ***What are the essential requirements of correlational statements?***

- Based on a specific, quantified relationship or association between one malleable measurement (like learner behavior or product/service usage) and one learner outcome of interest
- Refers to the specific study sample to which the statement applies
- Language expresses a directional statistical association within the same group

#### ***Illustrative examples of correlational statements***

- Among students using the product, learners who accessed more hints were more likely to earn a final grade of B or better
- Learners completing # additional homework problems were # times more likely to pass the course

### Comparative efficacy statements with caveats

---

#### ***Which research designs support comparative efficacy statements with caveats?***

Any research design that uses a treatment and comparison group with controls for baseline factors — although baseline equivalence of the two groups is not required — and at least one critical confound is known. Designs may include quasi-experimental studies where baseline equivalence is not achieved and where there is at least one known confounding factor.

#### ***What are the essential requirements of comparative efficacy statements with caveats?***

- Contains a specific, quantified measure of size (such as coefficient or effect size), significance, and direction of group difference based on analysis results described in the study
- Refers to the applicable treatment and comparison groups
- Does not meet one or more of the requirements of comparative efficacy statements (while still meeting the research design standards described on pages 22–24)
- Accompanied by a framing paragraph including caveats for the comparative efficacy statement requirements that were not met
- Language expresses directional statistical associations between two or more groups, along with at least one key caveat regarding rigor quality

#### ***Illustrative examples of comparative statements with caveats***

- When learners were given access to the product as an optional resource, those who used it scored on average # percentage points higher on final exams than those who did not use it

## Comparative efficacy statements

---

### ***Which research designs support comparative statements?***

Those that use a treatment and comparison group, with controls for baseline factors, although baseline equivalence of the two groups is not required. Designs may include quasi-experimental studies or randomized controlled trials where baseline equivalence is not achieved or where attrition patterns do not meet What Works Clearinghouse guidelines.

### ***What are the essential requirements of comparative statements?***

- Contains a specific, quantified measure of size (such as coefficient or effect size), significance, and direction of group difference based on analysis results described in the study
- Refers to the applicable treatment and comparison groups
- Based on measurements from a randomized controlled trial or quasi-experimental design with:
  - pre-/post-measurements (the pre-test measure must be a measure of achievement and must be correlated with the outcome measure, but need not be in the same domain)
  - a comparison group that was not exposed to the intervention
  - (for a randomized controlled trial) a comparison group that was determined at random
- Based on a design that avoids  $n=1$  perfect confounding
- Based on analysis that describes the extent to which the groups were similar or different and matched (if appropriate) on key baseline characteristics
- Based on analysis statistically controlling for key characteristics that are greater than 0.25 standard deviations different at baseline
- Language expresses directional statistical associations between two or more groups

### ***Illustrative examples of comparative statements***

- Learners using the product reported a better learning experience than learners who did not use the product, when controlling for prior achievement
- Learners using the product were #% more likely to progress to college level than learners who used a competing product, when controlling for prior achievement

## Causal efficacy statements

---

### ***Which research designs support causal statements?***

Those that use a treatment and comparison group, and demonstrate baseline equivalence of the two groups. Attrition must be within What Works Clearinghouse guidelines. Designs may include quasi-experimental approaches (propensity score matching, instrumental variables, regression discontinuity, fuzzy regression discontinuity) or randomized controlled trials.

### ***What are the essential requirements of causal statements?***

- Contains a specific, quantified measure of size (such as coefficient or effect size), significance, and direction of group difference based on analysis results described in the study
- Refers to the applicable treatment and comparison groups
- Based on measurements from a randomized controlled trial or quasi-experimental design with:
  - pre-/post-measurements
  - a comparison group that was not exposed to the intervention
  - (for a randomized controlled trial) a comparison group that was determined at random
- Based on a design that avoids:
  - $n=1$  perfect confounding
  - systematic differences between treatment and control on some observed covariate related to the outcome (such as if the treatment and control outcomes were measured in different cohorts or states)
  - confounding with a second intervention, unless the study intends to capture their joint efficacy (for example, if all treatment classrooms receive both MyLab Statistics and MyLab Mathematics, the treatment effect must be described as the effect of both products)
- Language expresses directional statistical associations between two or more groups, using language that refers to causal effects attributed to an intervention or treatment

If using a randomized controlled trial design:

- All assignments are determined at random
- The method of randomization is explicitly described (for example, student-level randomization or school-level randomization)
- Analytic sample excludes late joiners
- Reports on overall and differential attrition at student and, if applicable, cluster level
- Overall and differential attrition are compared to What Works Clearinghouse bounds (WWC Standards Handbook v.4, p.9–14) and authors acknowledge the potential for bias relative to WWC conservative and liberal attrition standards
- If attrition does not meet the conservative boundary, meets quasi-experimental design baseline equivalence standard

If using a quasi-experimental design:

- Baseline equivalence was achieved on prior achievement according to WWC guidelines (the baseline difference between treatment and control is within .25 SD and, if greater than .05, the measure is included as a covariate)
- Baseline equivalence was achieved for any other variables identified by the study as key characteristics (for example, student mobility, if student mobility is highly relevant to the context of the study)
- Other key confounding variables identified in the report are included as covariates in the analysis

### ***Illustrative examples of causal statements***

- Using the product positively influenced course pass rates relative to a different product
- Using the product boosted course pass rates by #% compared to students using a different product

To see the guidance document Pearson refers to when creating efficacy statements, see [appendix 1](#).

### **Research design standards**

---

Pearson developed a set of research quality criteria for guiding research design, and worked with SRI Education to codify these research quality criteria into tiered levels of evidence to reflect best practices in the field. Criteria for these tiered levels of evidence drew both from the tiered evidence levels in the USA's Every Student Succeeds Act (ESSA) and the 2006 American Educational Research Association's Standards for Reporting on Empirical Social Science Research.

Before deriving efficacy statements from any type of study, we review the research to make sure it meets these research quality criteria.

<b>Review category</b>	<b>Review questions</b>
Main findings	<p>Does the research clearly identify “main findings” for the reader?</p> <p>Did the main findings include an analysis of the full sample of interest? (This full sample may exclude students because of missing data.)</p> <p>Were overall outcome measures used in the main findings (for example, using a full assessment score and not using sub-scales)?</p> <p>Were the relevant outcomes for the main findings measured at one or more specific times appropriate for the study’s research questions?</p> <p>Were multiple comparison adjustments appropriately applied?</p>
Outcome measures	<p>Are the outcome measures used in the main findings of the study aligned to the Pearson’s learner outcomes? For example:</p> <ul style="list-style-type: none"> <li>• Achievement: assessment measuring students’ academic achievement</li> <li>• Timeliness and completion: students’ successful completion of, and/or on-time progression through, courses or grade levels</li> <li>• Progression: students’ successful completion of, and/or progression through, courses, grade levels, graduation rates, accessibility, and similar types of administrative measures</li> <li>• Employment, earnings, and other outcomes that are relevant to students’ educational goals</li> <li>• Social and behavioral competencies, skills, attitudes, and behaviors, including engagement and self-efficacy</li> <li>• Other outcomes that are believed to have positive effects on the above outcomes (for example, instructors’ self-efficacy can affect student outcomes, and cost effectiveness can improve student access)</li> </ul> <p>Do the outcome measures used in the main findings have face validity? That is, would experts and stakeholders agree that the outcomes appear to measure the constructs well?</p> <p>Are the outcome measures used in the main findings not overly aligned with the intervention?</p> <p>If the outcome measures are within Pearson’s control (for example, a Pearson assessment, usage of a Pearson product) a study-developed or administered measure, do the academic outcome measures used in the main findings meet the What Works Clearinghouse reliability standards for at least one of the following?</p> <ul style="list-style-type: none"> <li>• internal consistency of 0.50 or higher</li> <li>• temporal stability/test-retest reliability of 0.40 or higher</li> <li>• inter-rater reliability (such as percentage agreement, correlation, or kappa) of 0.50 or higher</li> </ul> <p>Do the main findings include analysis of at least one of the following academic outcome measures using the full study population?</p> <ul style="list-style-type: none"> <li>• an external/third-party test with established validity and reliability evidence</li> <li>• a widely used standardized assessment in a relevant domain (such as a state achievement test or college placement exam)</li> <li>• instructor-made course exams</li> <li>• course grades</li> </ul>



- Pearson's internal, summative platform measures — in which case it must meet:
  - internal consistency of 0.50 or higher
  - documentation of content-related validity evidence
  - items used on the outcome measure must be new to the students (in other words, test-retest validity threat must be controlled for)

For studies with multiple conditions, were outcome measures for main effects collected in the same manner for all conditions?

Research study sample	<p>Are the overall study sample and analytic samples relevant to the target population of the study (for example, do they draw from one of the key market segments identified for a Pearson product)?</p> <p>Do student-based analyses have analytic sample sizes of at least 100 students?</p> <p>Do teacher-based analyses include at least 66% of the relevant teachers? (For example, if the study includes a survey of teachers, was the survey response rate at least 66%?)</p>
Handling of missing data	<p>Assuming missing data are missing completely at random or missing at random, are missing data consistently handled using one or more of the following appropriate methods?</p> <ul style="list-style-type: none"> <li>• complete case analysis</li> <li>• regression imputation</li> <li>• full information maximum likelihood</li> <li>• non-response weighting</li> <li>• replacing missing data with a constant, combined with a missing data indicator</li> <li>• another approach justified as appropriate to the case and demonstrated within the study</li> </ul>
Handling of nested data	<p>Do the non-descriptive analyses need to account for nested data?</p> <p>Are nested data consistently handled using one or more of the following appropriate methods?</p> <ul style="list-style-type: none"> <li>• intraclass correlation coefficient calculated to be 0.05 or less and nesting therefore not modeled</li> <li>• multi-level modeling</li> <li>• clusters modeled using fixed effects (applied only if treatment varies within cluster)</li> <li>• clustering correction applied to standard errors</li> <li>• another approach justified as appropriate to the case and demonstrated within the study</li> </ul>
Covariates	<p>Do the non-descriptive student analyses include at least one exogenous student-level prior achievement covariate, which must be:</p> <ul style="list-style-type: none"> <li>• exogenous, that is, collected in such a way that student scores are independent of the study (such as selection into treatment or the effects of the treatment itself)</li> <li>• described to the reader along with its relevant domain (mathematics achievement, for example) and timing of collection</li> <li>• measured at baseline (including shortly after treatment begins)</li> <li>• reflective of achievement (such as cumulative GPA, standardized scores, pre-test performance, or diagnostic test score)</li> </ul>

## Audit of efficacy statements

Pearson has commissioned PwC to audit the efficacy statements identified in our Product Efficacy Reports. The PwC assurance report is included within each Product Efficacy Report.

The scope, approach and limitations of PwC's work are set out below.

### Professional standards applied and level of assurance

PwC performs a limited assurance engagement in accordance with International Standard on Assurance Engagements 3000 (Revised), *Assurance Engagements Other Than Audits and Reviews of Historical Financial Information*, issued by the International Auditing and Assurance Standards board.

A limited assurance engagement is substantially smaller in scope than a reasonable assurance engagement in relation to both the risk assessment procedures (including an understanding of internal control) and the procedures performed in response to the assessed risks.

### What is a material misstatement?

A material misstatement would be an efficacy statement that does not reflect the study design and quality of underlying research or the omission of key information from a relevant study.

### Work performed by PwC

PwC's audit focuses on:

- the process that generates the efficacy statements that appear in the Product Efficacy Report
- the integrity of the efficacy statements themselves

PwC's work includes the following procedures:

- Making enquiries of relevant Pearson management
- Evaluating the design of the Efficacy Reporting Framework including key structures, systems, processes and controls for managing, generating and reporting the efficacy statements
- Testing all the controls across the four processes of the Efficacy Reporting Framework
- Confirming that management reviews are performed over key sign-off artefacts in each process
- Performing substantive testing, on a sample basis, of the data that underpins the research studies and the resulting efficacy statements, and the controls over the completeness and accuracy of that data (supported by Pearson's internal audit team in those instances where student data is subject to confidentiality restrictions)
- Assessing the quality and conclusions of the underlying research studies
- Inspecting the statistical analysis to assess whether the efficacy statements are valid, supportable and consistent with the underlying research studies
- Independently re-performing the screening of relevant external public research studies and comparing to that done by Pearson
- Assessing the efficacy statements and underlying Technical Research Report(s) for consistency with the Efficacy Reporting Framework
- Reviewing the Product Efficacy Report and Technical Research Report(s) for alignment of research studies and efficacy statements

### Important limitations

- The assurance report is product-specific.
- Efficacy research studies reflect the implementation and use of a product in a particular context. It would not be appropriate to assume a product would always generate similar outcomes in the future.

## Appendix 1: Pearson efficacy statements guidance

Efficacy statements represent the main conclusions, headlines, or findings from an efficacy report that provide answers to the main research questions. Statements can be general (expressing a trend, relationship or effect in simple and broad terms) or specific (providing more detail). Decisions about whether efficacy statements should be general, specific, or both are at Pearson's discretion, but will generally be based on results of message testing and feedback from key audiences.

### General principles

- The shorter the better
- The more direct the better
- Use language that is accessible to someone without a PhD (or that means the same thing whether the reader has a PhD or not)

### Framing paragraph

Efficacy statements do not stand in isolation, but instead must be interpreted within the context of the study that generated the data. Wherever the efficacy statements appear, a short description of the study must also appear, either as a framing introductory paragraph to the efficacy statements (for general statements) or as a footnote (for specific statements). That description should include:

- The name of the institution(s) Pearson partnered with on the study (where appropriate)
- The title of the course(s)
- The year(s) of data collection
- The final analytic sample size
- The type of statistical model used and any covariates included (for example multiple regression model controlling for age, gender, and ACT scores)

*Note: The illustrative example efficacy statements presented here represent the types of wording that should be used to express different types of efficacy statements. These examples should not be construed as comprehensive or exhaustive. The wording of efficacy statements may vary slightly to improve readability, as long as a third-party validator agrees that these variations still preserve the original meaning and intent.*

<b>Efficacy statement type</b>	<b>Illustrative examples (general statements)</b>	<b>Illustrative examples (specific statements)</b>	<b>Requirements</b>
<b><i>Descriptive</i></b>	<ol style="list-style-type: none"> <li>Nearly all students reported that the product was a more efficient way to study compared to other books or products</li> <li>The average final exam score was in the A-range<sup>1</sup></li> </ol>	<ol style="list-style-type: none"> <li>#% of students said the product was a more efficient way to study compared to other books or products</li> <li>The average final exam score was #%</li> </ol>	<ul style="list-style-type: none"> <li>Identifies the specific sample to which the statement applies, if not the full study sample</li> <li>Explicitly states the outcome measure used</li> <li>Does not imply a quantitative correlation or statistical association</li> </ul>
<b><i>Correlational</i></b>	<ol style="list-style-type: none"> <li>On average, students earning higher quiz scores in the product also tended to earn higher unit exam scores</li> <li>Attempting more homework questions using the product is related to higher final exam scores</li> </ol>	<ol style="list-style-type: none"> <li>Each additional ten percentage points students scored on quizzes in the product were associated with an increase of # percentage points on unit exams</li> <li>A #% increase in the percentage of unique homework questions attempted using the product is associated with a #% increase in final exam scores</li> </ol>	<ul style="list-style-type: none"> <li>Identifies the specific sample to which the statement applies, if not the full study sample</li> <li>Clearly states where a null result was observed for a main study finding (<math>p &gt; .05</math>)</li> <li>Explicitly states the outcome measure used</li> <li>Indicates the direction of the relationship</li> <li>Specific statements express the magnitude of the effect in practical terms</li> <li>Does not include details about the type of model or covariates included</li> <li>Avoids causal language</li> <li>Clearly specifies whether referring to the percent metric or percentage point metric</li> </ul>
<b><i>Comparative with caveats</i></b>	<ol style="list-style-type: none"> <li>Students using the product in spring 2019 outperformed students using other products in terms of the percentage scoring proficient on state reading assessments</li> </ol>	<ol style="list-style-type: none"> <li>When college freshmen were given access to the product as an optional resource, those who used it scored on average # percentage points higher on final exams than those who did not</li> </ol>	<ul style="list-style-type: none"> <li>As with comparative statements, plus...</li> <li>Includes (one of) the most relevant comparative rigor standards that were not met, making clear that the unmet rigor standard is related to the efficacy statement</li> <li>Framing paragraph includes all comparative rigor standards that were not met</li> </ul>

<sup>1</sup> General descriptive statements are uncommon. Where such statements are used, further information should be included within the framing paragraph, such as details about the evidence that supports the statement.

Efficacy statement type	Illustrative examples (general statements)	Illustrative examples (specific statements)	Requirements
<b>Comparative</b>	<ol style="list-style-type: none"> <li>1. There was no difference in the percentage of students deemed proficient on state math and reading assessments between students using the product and those using a different product</li> <li>2. Student cohorts using the product outperformed students using a different product in terms of the percentage scoring proficient on state reading assessments</li> </ol>	<ol style="list-style-type: none"> <li>1. Students using the product were almost twice as likely to pass the course than students who used a different product</li> <li>2. The percentage of students deemed proficient on state reading assessments was # points higher for cohorts using the product than for cohorts using a different product</li> </ol>	<ul style="list-style-type: none"> <li>• Clearly states where a null result was observed for a main study finding (<math>p &gt; .05</math>)</li> <li>• Specifies the treatment and comparison groups</li> <li>• Indicates direction of any differences between groups</li> <li>• Explicitly states the outcome measure used</li> <li>• Specific statements express the magnitude of group differences in practical terms</li> <li>• Does not include details about the type of model or covariates included; can mention that the treatment and control groups are similar if the statement also includes the covariates they are similar on</li> <li>• Avoids causal language</li> </ul>
<b>Causal</b>	<ol style="list-style-type: none"> <li>1. Using the product positively influenced course pass rates relative to a different product</li> </ol>	<ol style="list-style-type: none"> <li>1. Using the product boosted course pass rates by #% compared to students using a different product</li> </ol>	<ul style="list-style-type: none"> <li>• Clearly states where a null result was observed for a main study finding (<math>p &gt; .05</math>)</li> <li>• Specifies the treatment and comparison groups</li> <li>• Indicates direction of any differences between groups</li> <li>• Explicitly states the outcome measure used</li> <li>• Specific statements express the magnitude of group differences in practical terms</li> <li>• Does not include details about the type of model or covariates included</li> </ul>



Pearson