Analysis of Patterns in Time for Evaluating Effectiveness of First Principles of Instruction

Theodore W. Frick
Professor Emeritus
Department of Instructional Systems Technology
School of Education, Indiana University Bloomington, USA

Rodney D. Myers
Instructional Consultant
School of Education, Indiana University Bloomington, USA

Cesur Dagli
Research Analyst
Office of Analytics & Institutional Effectiveness
Virginia Polytechnic Institute and State University, USA

Featured Research

This version of the article has been accepted for publication, after peer review (when applicable) but is not the Version of Record and does not reflect post-acceptance improvements, or any corrections. The Version of Record is available online at: http://dx.doi.org/10.1007/s11423-021-10077-6. Use of this Accepted Version is subject to the publisher’s Accepted Manuscript terms of use https://www.springernature.com/gp/open-research/policies/accepted-manuscript-terms.
Analysis of Patterns in Time for Evaluating Effectiveness of
First Principles of Instruction

Abstract

In this naturalistic design-research study, we tracked 172,417 learning journeys of students who were interacting with an online resource, the Indiana University Plagiarism Tutorials and Tests (IPTAT) at https://plagiarism.iu.edu. IPTAT was designed using First Principles of Instruction (FPI; Merrill, 2002, 2013, 2020). Students who used IPTAT were mostly from university and advanced high school courses taught in 186 countries and territories. Instructors expected their students to pass one of trillions of difficult Certification Tests (CT) provided through IPTAT. Each CT assessed student ability to classify samples of writing as word-for-word plagiarism, paraphrasing plagiarism, or no plagiarism—when given original source materials. In 51,646 successful learning journeys, students who were initially nonmasters and who later achieved mastery had viewed, on average, 89 IPTAT tutorial webpages designed with FPI. In the 23,307 unsuccessful learning journeys, students who were nonmasters and who had not (yet) achieved mastery had viewed an average of 52 tutorial webpages designed with FPI. Analysis of Patterns in Time (Frick, 1990) and Bayesian analysis revealed that students were nearly 4 times more likely to pass a CT when they selected one or more parts of IPTAT instruction designed with FPI. These results support the instrumental value of First Principles of Instruction for design of online learning in a massive, open, online course (MOOC). These findings further demonstrate the value of an innovative approach to modern learning analytics, Analysis of Patterns in Time (APT), when coupled with Bayesian reasoning.
**Keywords**

Online learning; learning journeys; innovative learning analytics; MOOC; First Principles of Instruction; Analysis of Patterns in Time; instructional effectiveness; recognizing plagiarism; certification tests; mastery learning; web-based instruction.
Analysis of Patterns in Time for Evaluating Effectiveness of
First Principles of Instruction

Theoretical Framework

Michael Scriven (1967) has been often cited for first introducing the terms formative versus summative methods when evaluating curriculum in education. Broadly conceived, these methods of evaluation are distinguished by their purpose. Formative evaluation is used to improve something during its development, whereas summative evaluation is used to determine its merit or worth when that development is completed. Worthen and Sanders (1987) also made this distinction, emphasizing that formative evaluation entails asking questions such as “What is working?”, “What needs to be improved?”, and “How can it be improved?” (p. 36). Reigeluth and Frick (1999) recommended ways that formative research methodology can be used to evaluate instructional-design theories. They stated:

The underlying logic of formative research … is that if you create an accurate application of an instructional-design theory (or model), then any weaknesses that are found in the application may reflect weaknesses in the theory, and any improvements identified for the application may reflect ways to improve the theory. (p. 636)

The broader issue is praxiology in educology, according to Steiner (1988). Disciplined inquiry about education, if adequate, results in qualitative, quantitative, and performative educology (Steiner, 1988). Educology is in essence “recorded signs of knowing about education” (Frick, 2021, p. 28).
Quantitative knowledge is comprised of scientific, praxiological, and philosophical theories which have been verified by appropriate research methods for disciplined inquiry. These theories are about universals. Universals are not limited to time or place (Steiner, 1988).

A praxiological theory is comprised of recorded signs about means-ends relationships that are universal. Such means-ends relationships are not limited to time or place. Means that successfully lead to achieving ends have instrumental value—i.e., the means are good for the ends (Steiner, 1988). Such means are effective. Instructional-design (ID) theories are, therefore, a type of praxiological educology. ID theories specify ways of guiding student learning (means-ends relationships) that can have broad generalizability (e.g., see Frick & Reigeluth, 1992).

David Merrill posited what he termed First Principles of Instruction (FPI; 2002, 2013, 2020). First Principles can be considered as an instructional-design theory, and hence also are part of praxiological theory in educology. Merrill (2020) defined a principle as a relationship between learning outcomes and instructional strategies that is always true under appropriate conditions regardless of the methods or models used to implement it. Rather than methods or models of instruction themselves, principles are the relationships that may underlie any model or method. (p. 2)

**Overview of this Study**

This study is best characterized as design research, with the goal of adding to praxiological knowledge of education (DBRC, 2003). It is not an experimental study—there is no random assignment of participants to treatment and control groups. In fact, there is no uniform “treatment” because the participants were free to choose parts of instruction in which to
engage within the online Indiana University Plagiarism Tutorials and Tests (IPTAT). Thus, this is a naturalistic study to do design research.

Praxiological knowledge characterizes means-ends relationships—i.e., what works to achieve goals that are sought. A goal is an end; and means have instrumental value if they are likely to result in achieving that goal. In the case of IPTAT, the means are numerous components of instruction through which students can choose to navigate during their learning journeys via their Web browsers. The end (goal) is normally to pass a Certification Test, which is an indicator of student mastery with respect to their ability to recognize plagiarism—specifically to discriminate correctly word-for-word plagiarism from paraphrasing plagiarism from non-plagiarism when provided with a writing sample and original source material.

In contrast, scientific knowledge of education is about universals and their non-axiological relationships (Steiner, 1988). Scientific knowledge characterizes what is—and the primary goal of scientific research is to determine truth of claims about generalizable relationships, not to assess instrumental value of means to reach ends (Steiner, 1988). Relationships identified in scientific theories are not characterized by their value with respect to human endeavors—either instrumental or intrinsic.

When humans try to do something, we act intentionally. Education should be an intentional human endeavor. Teachers attempt to guide student learning, and students try to learn (Steiner, 1988; Frick, 2021). Means of guidance are instructional strategies and methods, and the desired end is successful student learning achievement.

On the other hand, Newton’s laws of gravity are part of a scientific theory. In theoretical physics, there is a generalizable relationship between force, mass of a body, and its acceleration \((F = ma)\). In human learning, a scientific relationship is that learner experience is associated
with formation of new mental structures. Network neuroscience is investigating how experience physically changes in the brain’s structure by increasing the complexity of neuronal connections (Bertolero & Bassett, 2019; Eagleman, 2020). These are examples of non-axiological relations in physics and human learning, respectively. These relations are not about intrinsic or instrumental value. Axiology concerns value (Steiner, 1988).

In further contrast, philosophical knowledge is not about what is or what works, but rather what is worthwhile in itself—i.e., what has intrinsic value. What is worthwhile may not currently exist, and what exists is not necessarily worthwhile. For example, it is a fact that some humans murder other human beings. However, just because murder exists does not imply that it is worthwhile. Murdering of humans cannot be rationally justified on moral grounds, if the categorical imperative is used as a criterion for judging ethical behavior (Kant, 1785).

Using a gun to commit murder has instrumental value. Guns are often an effective means to kill people. However, the end cannot be justified as having intrinsic value when Kant’s categorical imperative is used for rationally determining justice. The means do not justify the end.

In philosophical educology, Steiner (1981) argues that education should lead students to become rational and thus free. She makes the case that student rationality is a worthwhile goal for educators to pursue. For further details, see Steiner (1981, 1988, 2009) and the educology website: https://educology.iu.edu (n.d.).

To summarize, in the present study the means are various instructional components of IPTAT, and the end is for students to pass a Certification Test on classifying plagiarism and non-plagiarism. We are concerned about the instrumental value of those means.
Frick et al. (2022) provided the example of the Oregon Trail as a metaphor to illustrate means-end relationships. The Oregon Trail was a major route from what is now Kansas City, Missouri to Portland, Oregon in the U.S. In the mid-1800s, the means by which people undertook this journey was by riding in covered wagons pulled by oxen, by riding horses, by walking, and often by combinations of these means of conveyance. In modern times, different means are available. These include riding a bicycle, driving an automobile, or flying as passengers on commercial airplanes.

The successful end of the journey is to arrive at the destination, Portland, Oregon, while surviving the trip. In the mid-1800s, not everyone was successful. Disease or death prevented some travelers from achieving their goal. Others chose to change their goal and settle along the route, never reaching Portland. Those were unsuccessful journeys with respect to reaching their original goal. Even in modern times, some journeys are unsuccessful—bicycles and automobiles can break down, airplanes can crash, and illness or death can prevent travelers from reaching their destination. Even though means can often be effective in reaching ends, sometimes they are not. Moreover, this does not imply that means cause the ends, rather those means have instrumental value.

Instrumental value of means can further be examined beyond their success rate. For example, Frick et al. (2022) considered cost and duration of Oregon Trail trips, diseases contracted, deaths, and reasons for dropping out. Reigeluth and Frick (1999) also described these additional factors as part of an instructional method’s appeal to students and teachers—i.e., which means they prefer.
Description of this Design Study

Merrill’s First Principles of Instruction were applied in 2015 when redesigning the online Indiana University Plagiarism Tutorials and Tests (IPTAT). The praxiological study described in this article is a form of **summative evaluation** (Scriven, 1967; Worthen & Sanders, 1987). We used a method to verify praxiological theory called *Analysis of Patterns in Time* (APT; Frick, 1990; Frick et al., 2022; Frick & Reigeluth, 1992; Myers & Frick, 2015).

While APT has been used in previous studies over the past four decades (see Chapters 5-8, Frick et al., 2022), what is innovative in this study is the use of Google Analytics 4 (GA4) to track individual student interaction with IPTAT. GA4 is a relatively new analytic tool, first made available in mid-October 2020. To our knowledge there have been no previous studies which have used GA4 to do APT of big data, as we now illustrate below in the present study. Frick et al. (2022) had used Google’s Universal Analytics (UA) in their Big Study in 2019 and 2020, also a new way to do learning analytics with APT.

GA4 created the temporal maps needed for APT—that is, big data collected on 172,417 learning journeys through IPTAT during early 2021. See Figure 1 below for an example of a temporal map. GA4 segmenting and matching tools were subsequently applied to count event patterns within those temporal maps which indicated student experience of First Principles of Instruction and their learning outcomes. Results from GA4 queries were then exported to a Microsoft Excel spreadsheet to perform further computations needed for APT and Bayesian reasoning.

It is worth noting that other analytic approaches could have been employed. For example, a linear models approach (LMA) could have been used. LMA includes multiple and logistical regression, ANOVA, MANOVA, ANCOVA, discriminate analysis, time-series
analysis, factor analysis, and the like. These are linear, additive models based on an algebraic function for a line in a Cartesian coordinate system (see Kirk, 1999, 2013; Tabachnick & Fidell, 2018). However, Frick (1983, 1990) clearly demonstrated empirically and proved mathematically that stochastic educational relationships cannot be verified as precisely by the LMA when compared to APT. Myers and Frick (2015) characterized the difference as follows: The LMA relates measures, whereas APT measures relations. This is not wordplay, but a fundamental difference in approach to measurement and analysis of stochastic relations.

When we use LMA, we measure things independently as separate variables, and then we use an appropriate statistical linear model to determine relationships among variables. In the LMA, the strength of a relationship is often characterized by percent of variance accounted for in the independent variable by one or more dependent variables (e.g., see Kirk, 1999, 2013; Tabachnick & Fidell, 2018). The LMA is appropriate for verification of deterministic relations among separate variables or factors (Frick, 1983; Myers & Frick, 2015).

On the other hand, in APT we observe and record event occurrences in temporal maps for each unique case. Temporal maps are the data source needed to do subsequent APT queries. APT queries do the segmenting and matching in order to count event patterns within temporal maps. These APT queries result in estimates of the likelihoods of patterns in terms of their probability of occurrence. In APT, we can predict the probability of a consequent event, given an antecedent event or set of events. Noteworthy is that APT is fundamentally grounded in set theory and probability theory in mathematics. When these APT assumptions are met, APT is appropriate for estimating likelihoods of stochastic relations (temporal patterns). APT is best used when theory guides disciplined inquiry. APT does not prescribe which patterns to look for; rather theory should.
In both the LMA and APT we normally attempt to make predictions. But different assumptions are made. Frick et al. (2022) described in more detail the differences in approach. The main difference is that APT can discriminate patterns which the LMA cannot—by virtue of the fundamental difference in approach to measuring relationships. This was proven mathematically and demonstrated empirically (see Frick, 1983, pp. 72-74). APT directly measures relations as temporal patterns; the LMA statistically relates separate measures of variables through linear, additive models (i.e., mathematical functions). The LMA cannot be theoretically reduced to APT; whereas results from APT can be further analyzed by LMA when appropriate (Frick, 1983). When variables are measured separately in the LMA, patterns can be obfuscated—particularly when relations are stochastic (i.e., not deterministic).

**Primary Research Question**

APT methods allowed us to address the primary research question:

What is the likelihood of achieving mastery when students select IPTAT webpages designed with First Principles of Instruction (FPI)?

Student learning achievement in IPTAT (i.e., student mastery) was measured by well-established and reliable Certification Tests (CTs, Frick & Dagli, 2016). Students were considered successful if they answered at least 9 questions out of 10 randomly selected questions on a CT. Test questions were randomly selected from large pools, resulting in trillions of unique tests. Merrill (2020) had hypothesized that:

- when a given instructional program or practice implements one or more of these First Principles, there will be an increase in learning and performance.

Obviously, the support for this hypothesis can only come from evaluation.
studies for a given instructional product or research comparing the use and misuse of these principles. (p. 3)

Indeed, the IPTAT is “a given instructional product” which, in this case, happens to be a massive, open, online course (MOOC; Frick & Dagli, 2016); and the present design-research study is an evaluation of IPTAT’s instrumental value, based on how it is used by students. When students are initially nonmasters and they later become masters, this is an increase in their learning and performance. In IPTAT, student mastery is indicated by answering 9 or 10 questions correctly on a CT (i.e., they pass the test). Nonmastery is indicated by answering fewer than 9 questions correctly (they do not pass the test). Students who move from a state of nonmastery at time 1 to a state of mastery at time 2 have achieved the learning objective. They now are able to do something they were unable to do previously. This indicates an “increase in learning and performance” from time 1 to time 2, which is how we have investigated Merrill’s (2020) claim.

Methods

Redesign of IPTAT in 2015

Frick et al. (2018) described 14 years of IPTAT development and use historically. While originally designed for students in the Instructional Systems Technology program at Indiana University, IPTAT has subsequently been adopted by many instructors worldwide. These instructors want their students to avoid committing plagiarism, and typically they expect their students to pass an IPTAT Certification Test as a course requirement. Frick et al. (2022) indicated that IPTAT had been accessed over 125 million times since its inception, and that from 2016 through 2020 nearly 750,000 students had passed one of trillions of IPTAT’s difficult Certification Tests.
A team had significantly redesigned IPTAT in 2015 by applying First Principles of Instruction (Merrill, 2002, 2013, 2020). Merrill’s First Principles of Instruction as applied to IPTAT included:

- **Authentic problems** or tasks for students to do, arranged from simple to complex (e.g., [https://plagiarism.iu.edu/tutorials/index.html](https://plagiarism.iu.edu/tutorials/index.html));
- **Activation** of student learning by helping students connect new learning with what they already know or believe (e.g., [https://plagiarism.iu.edu/tutorials/task1/activation.html](https://plagiarism.iu.edu/tutorials/task1/activation.html));
- **Demonstration** of what is to be learned, by showing a variety of examples (e.g., [https://plagiarism.iu.edu/tutorials/task1/demonstration.html](https://plagiarism.iu.edu/tutorials/task1/demonstration.html));
- **Application** of what is being learned, so students can try themselves and feedback is provided (e.g., [https://plagiarism.iu.edu/practiceTest.php?task=1&amp;item=1](https://plagiarism.iu.edu/practiceTest.php?task=1&amp;item=1)); and
- **Integration** of what has been learned into students’ own lives (e.g., [https://plagiarism.iu.edu/tutorials/task1/integration.html](https://plagiarism.iu.edu/tutorials/task1/integration.html)).

Examples of application of First Principles are illustrated by hyperlinks to webpages above. Since this design has been described in more detail elsewhere, readers are referred to Frick et al. (2018, 2022). After IU technology services had discontinued their in-house Web statistics, the design team incorporated Google Analytics in the new version of IPTAT, which went live on January 2, 2016. At that time, they did not know that they could do APT using Google Analytics. They wanted to know how IPTAT was being used by students, and how well it was working. They did not know that the present study using GA4 could even be done.

**Analysis of Patterns in Time**

The primary research method we used to evaluate the effectiveness of First Principles of Instruction in IPTAT was Analysis of Patterns in Time (Frick, 1990; Frick et al., 2022; Myers &
Frick, 2015). APT has been used in many past studies (see Frick et al., 2022). However, to our knowledge, Google Analytics 4 (GA4) has not been previously used to do APT. We ourselves only discovered this about two months after GA4 was officially released by Google. We describe in detail below how we subsequently leveraged GA4 to do APT when supplemented by spreadsheet calculations.

Frick et al. (2022) used the metaphor of the Oregon Trail, comparing how early settlers followed it by walking and riding in covered wagons with how modern-day transportation systems can be used to make this trip. They introduced the concept of learning journeys and briefly summarized the limitations of traditional quantitative and qualitative approaches. They described APT as an alternative, which instead uses temporal maps as the primary data collection source. APT thus allows researchers to document what happens during learning journeys. If enough learning journeys are sampled, researchers can make predictions about patterns of student success and failure that are associated with various instructional strategies.

The important discovery in 2020 by Frick et al. (2022) was that Google Analytics has been implementing many ideas from the original APT, which had been invented decades earlier (Frick, 1983, 1990). GA tracking constructs APT temporal maps, and if used creatively, GA analysis tools can subsequently do segmenting and matching within temporal maps, resulting in counts of event occurrences that have been previously tracked on user interaction with a website such as IPTAT. Results from GA reports can then be transferred to a spreadsheet, where cell formulae are created to do further APT computations needed for forming likelihoods and odds ratios.
Google Analytics 4 for Collecting Big Data

GA4 sessions are indeed temporal maps as described by Frick et al. (2008, 2022) and Myers and Frick (2015). A GA4 session is defined by user interaction with a website with no more than 30 minutes elapsing between pageviews. We consider a student learning journey through IPTAT to consist of one or more GA4 sessions (APT temporal maps). GA4 allowed us to subsequently segment those student learning journeys based on their navigation through IPTAT and whether or not they passed Certification Tests (CTs). GA4 reports refer to ‘Active Users’, which are equivalent to IPTAT student learning journeys in this study. IPTAT also independently stored records of CT results, which served to triangulate measures of student success identified in GA4 sessions.

Using GA4 to Carry Out APT Queries

Caveat: It took us some time to discover how to adapt GA4 in order to do Analysis of Patterns in Time. We were breaking new ground and not even sure we could do APT with GA4 when we started. Once we better understood how GA4 tracks and counts events, as well as how it identifies clients (Active Users), then we were able to proceed as described below. While GA4 by itself cannot do all of APT, it can greatly facilitate the counting process. Additional APT calculations such as means, likelihoods, and odds ratios can subsequently be calculated with a spreadsheet.

We hope that our descriptions below will help guide others to use this approach to praxiological research. We provide this detail so that other researchers can replicate our methods. To our knowledge, there currently are no textbooks or courses that teach researchers how to do this with GA4. For this reason, Frick et al. (2021) had further created video
demonstrations to illustrate in greater detail how Google Analytics can be used to do APT, now available at https://plagiarism.iu.edu/apt/demo/.

**Setting up Website Tracking for GA4.** IPTAT users were tracked since 2016 via Universal Analytics (UA), an earlier version of Google Analytics. Starting in early December 2020, we connected the existing UA tracking system to the new GA4 tracking system, and then enabled new GA4 tracking records.

When initially setting up Google Analytics, a snippet of JavaScript code was provided that contained our unique website ID, which the design team had inserted into the HTML templates for webpages to be tracked. Whenever anyone accesses a particular webpage, their browser executes the JavaScript code when the page is displayed. This code sends information to Google’s tracking system which includes the hashed client-ID of the user, webpage URL (path), HTML title of the page, IP address of the client device, and the current date and time. Specific website users remain anonymous, since their device’s IP address and client-ID are encrypted in GA tracking records to help protect user privacy.

GA normally can determine when the same device accesses a different webpage at a later time through use of browser cookies stored on that device. If users disable cookies, clear them from their browsers, or use browsers or settings that prevent tracking, then GA tracking methods are thwarted. This may be viewed as a limitation of GA, since users with privacy concerns can prevent tracking. This issue is addressed in the IPTAT privacy policy (https://plagiarism.iu.edu/privacy.html and https://plagiarism.iu.edu/protect.html). When users register for IPTAT, permission to use their data only in aggregate form is requested, such that they cannot be identified individually in published reports such as this research article.
Once tracking of a website is enabled, GA reporting tools can be used to analyze what users have done on the website. Google provides authentication methods for GA administration normally by using one’s Gmail account. This further restricts who can access the tracking data on a particular website.

When the design team initially built the IPTAT website at https://plagiarism.iu.edu they created page names that corresponded to various First Principles of Instruction and other important activities (e.g., /activation*, /demonstration*, /masteryTest*, /practice*, /plagiarismTestUG*, /mail*, etc.). Note that the asterisks (*) used here are wildcards for variations of webpage filenames. For example, any filename that contained ‘/activation’ corresponded to use of the FPI activation principle in IPTAT design.

The design team originally intended this file naming convention for their own benefit as instructional developers of the website. Fortuitously, this also simplified subsequent APT queries with GA4 tools for us to specify segments and matching conditions. For example, all FPI activation events can be identified by matching webpage filenames that contain the string, ‘/activation’. Or we can determine from GA4 that whenever a ‘/mail*’ webpage was accessed, this meant that a student had just passed a Certification Test (CT) and clicked the button to email it to themselves. Note that this particular webpage can only be accessed immediately after a student has passed a CT.

Creating New GA4 Events and Conversions.

GA4 tracks by default events that include pageviews, clicks, scrolling, and client session starts. To do APT, we needed to create new events according to pageviews of First Principles of Instruction. For example, the matching condition for an Activation event was whenever there was a ‘page_view’ event and the ‘page_location’ parameter contained “/activation”. All of the
IPTAT webpages that we specifically designed using the FPI Activation principle contain this string in their filename paths (e.g., https://plagiarism.iu.edu/tutorials/task4/activation2.html).

In early December 2020, we defined similar matching conditions accordingly for other FPI events. Once these new FPI events were created, they were further marked as GA4 conversions (i.e., FPI goals achieved by IPTAT users). In essence, by creating new events in GA4, we further configured it to classify various FPI webpages by appropriate categories (Frick, 1990; Myers & Frick, 2015). Note further that, in GA4 terminology, these conversions are goals. However, in the context of educational praxiology, the Web page activities for guiding of student learning are considered *means*, and the *end* is the student goal of passing a CT. While this may be confusing, we are retaining GA4 terminology here so that readers will understand how they can adapt GA4 to do APT. GA4 has its own lexicon and definitions of terms, as well as does APT (Frick, 1983, pp. 100-104).

This was a very important initial step. We initially configured GA4 this way because we later intended to do Analysis of Patterns in Time. This method of designing naturalistic research is an innovative approach via GA4, although APT itself is a research method that has been around for decades (Frick, 1990).

[Insert Figure 1 about here]

GA4 User Explorer View of Individual Temporal Map. Figure 1 illustrates an excerpt from a single learning journey through IPTAT, captured on March 25, 2021, from 6:30 to 7:03 p.m. Eastern Standard Time. The bottom section of the figure is part of the temporal map, listed in reverse chronological order for this user with client ID 187389469.1616711619. For example, an FPI Application event was last recorded at 7:03:39 p.m., and we can see in the row below that it was also coded as a ‘page_view’ event. The Application event was flagged as a ‘conversion’
event, whereas the ‘page_view’ is a GA4 default tracking event, which was also marked as a ‘conversion’. Earlier, at 7:03:34 p.m. a ‘scroll’ event had occurred. At 7:02:55 p.m. an FPI Demonstration event was recorded. And so forth. Figure 1 shows only 8 out of 70 total events for this user session. Note that every ‘conversion’ is also coded as a ‘page_view’, but not all ‘page_views’ are FPI ‘conversions’. Note further that these are not unique ‘page_views’. Every time the user’s device displays a webpage, it is coded as a new event with its corresponding time stamp.

**Figure 1**

*Excerpt of Individual Temporal Map and Associated Data for User Client ID: 187389469.1616711619*

In Figure 1, the temporal order is reversed chronologically when viewing from top to bottom. We can also see that this user’s client device was located in Grand Bay in the U.S. Summary data on this session is shown above the temporal map. Total user engagement during
this session was 33 minutes and 16 seconds. There were 19 Application events, 5 Demonstration events, 3 Activation events, and others not shown in Figure 1.

This excerpt of a temporal map is part of one student’s learning journey; and there were 172,417 such learning journeys in this study. In other words, these are big data on IPTAT usage from Jan. 1 to March 25, 2021. Further summative data are reported in figures and tables below.

[Insert Figures 2 – 8 and Tables 1 – 4 about here]

**Figure 2**

*GA4 Explore—Segment Overlap Analysis: Venn Diagram of IPTAT Active Users and Test Evaluations (the Subset Who Took Two or More Certification Tests and Received Test Feedback)*

![Segment Overlap Analysis Diagram](image-url)

<table>
<thead>
<tr>
<th>Segment set</th>
<th>Active users</th>
<th>Conversions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Totals*</td>
<td>173,936</td>
<td>16,547,092</td>
</tr>
<tr>
<td></td>
<td>100% of total</td>
<td>100% of total</td>
</tr>
<tr>
<td>1 IPTAT Active Users</td>
<td>172,417</td>
<td>16,553,187</td>
</tr>
<tr>
<td>2 IPTAT Active Users ONLY</td>
<td>97,029</td>
<td>1,642,404</td>
</tr>
<tr>
<td>3 IPTAT Active Users + Test Evaluations</td>
<td>75,087</td>
<td>14,910,783</td>
</tr>
</tbody>
</table>
**Figure 3**

*GA4 Explore—Segment Overlap Analysis: Venn Diagram of Two Segments: Test Evaluations x x Achievers, with Nonmasters Highlighted (Test Evaluations ONLY)*

![Figure 3](image)

**Figure 4**

*GA4 Explore—Segment Overlap Analysis: Venn Diagram of Three Segments: Test Evaluations x Try any FPI x Achievers*

![Figure 4](image)
Figure 5

GA4 Explore—Segment Overlap Analysis: Venn Diagram of Three Segments: Test Evaluations Highlighted

![Venn Diagram of Test Evaluations and Achievers](image1)

Figure 6

GA4 Explore—Segment Overlap Analysis: Venn Diagram of Three Segments: Achievers Highlighted

![Venn Diagram of Test Evaluations and Achievers](image2)
Figure 7

GA4 Explore—Segment Overlap Analysis: Venn Diagram of Three Segments: Try any FPI Highlighted

Figure 8

GA4 Explore—Segment Overlap Analysis: Venn Diagram of Three Segments: Intersection of IPTAT Active Users Who Received Test Evaluations and Who Tried any First Principle of Instruction and Who Subsequently Passed a Certification Test (Achievers)
Table 1

*Initial Segment Types Identified through the GA4 Explore Tool: Segment Overlap Analysis of IPTAT Active Users*

<table>
<thead>
<tr>
<th>Segment Type</th>
<th>Active Users</th>
<th>Minutes of User Engagement</th>
<th>Minutes of Engagement per User</th>
<th>Pageviews</th>
<th>Pageviews per User</th>
</tr>
</thead>
<tbody>
<tr>
<td>IPTAT Users</td>
<td>172,417</td>
<td>5,944,940</td>
<td>34.5</td>
<td>8,276,594</td>
<td>48.0</td>
</tr>
<tr>
<td>IPTAT Active Users ONLY (Dabblers)*</td>
<td>97,029</td>
<td>327,116</td>
<td>3.4</td>
<td>821,202</td>
<td>8.5</td>
</tr>
<tr>
<td>IPTAT Active Users AND Test Evaluations</td>
<td>75,087</td>
<td>5,617,823</td>
<td>74.8</td>
<td>7,455,392</td>
<td>99.3</td>
</tr>
<tr>
<td>Test Evaluations AND Achievers (Achievers)</td>
<td>51,646</td>
<td>4,078,607</td>
<td>79.0</td>
<td>5,606,085</td>
<td>108.5</td>
</tr>
<tr>
<td>Test Evaluations ONLY (Nonmasters)</td>
<td>23,307</td>
<td>1,539,216</td>
<td>66.0</td>
<td>1,849,307</td>
<td>79.3</td>
</tr>
<tr>
<td>Achievers ONLY (Already Achievers)*</td>
<td>10</td>
<td>105</td>
<td>10.5</td>
<td>102</td>
<td>10.2</td>
</tr>
</tbody>
</table>

*‘Dabblers’ took less than 2 CTs and spent little time on the IPTAT website (3.4 min. on average). ‘Already Achievers’ passed a Certification Test on their first try (a rare event), and hence were not part of the Test Evaluations Segment—that is, those who took two or more CTs and failed at least one of them. Already Achievers and Dabblers were excluded in subsequent analyses.*
Table 2

Further Segment Types when First Principles of Instruction (FPI) is Added to the Mix: Segment Overlap of IPTAT Active Users: Test Evaluations x Try any FPI x Achievers

<table>
<thead>
<tr>
<th>GA4 Segment Type</th>
<th>Active Users</th>
<th>Minutes of User Engagement</th>
<th>Minutes of Engagement per User</th>
<th>Pageviews</th>
<th>Pageviews per User</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Test Evaluations</td>
<td>75,087</td>
<td>5,617,823</td>
<td>74.8</td>
<td>7,455,392</td>
<td>99.3</td>
</tr>
<tr>
<td>2. Try any FPI</td>
<td>73,659</td>
<td>5,409,684</td>
<td>73.4</td>
<td>7,782,421</td>
<td>105.7</td>
</tr>
<tr>
<td>3. Test Evaluations AND Try any FPI</td>
<td>58,078</td>
<td>5,156,235</td>
<td>88.8</td>
<td>7,107,175</td>
<td>122.4</td>
</tr>
<tr>
<td>4. Test Evaluations AND Achievers</td>
<td>51,646</td>
<td>4,078,607</td>
<td>79.0</td>
<td>5,606,085</td>
<td>108.5</td>
</tr>
<tr>
<td>5. Test Evaluations AND Achievers AND Try any FPI</td>
<td>42,046</td>
<td>3,801,331</td>
<td>90.4</td>
<td>5,402,396</td>
<td>128.5</td>
</tr>
<tr>
<td>6. Test Evaluations AND Achievers AND NOT Try any FPI</td>
<td>9,600</td>
<td>277,276</td>
<td>28.9</td>
<td>203,689</td>
<td>21.2</td>
</tr>
<tr>
<td>7. Test Evaluations ONLY AND NOT Achievers</td>
<td>23,307</td>
<td>1,539,216</td>
<td>66.0</td>
<td>1,849,307</td>
<td>79.3</td>
</tr>
<tr>
<td>8. NOT Achievers AND Try any FPI</td>
<td>16,032</td>
<td>1,354,904</td>
<td>84.5</td>
<td>1,704,780</td>
<td>106.3</td>
</tr>
<tr>
<td>9. NOT Achievers AND NOT Try any FPI</td>
<td>7,448</td>
<td>184,313</td>
<td>24.7</td>
<td>144,528</td>
<td>19.4</td>
</tr>
</tbody>
</table>

Key:
1. Took two or more Certification Tests and got feedback, at least one of which was failed
2. Viewed one or more FPI-designed webpages (tried any FPI)
3. Traditionalist: Got test evaluations and tried any FPI
4. Achiever: Failed one or more tests, then passed
5. Traditionalist Achiever: Tried any FPI and failed one or more tests, then passed
6. Minimalist Achiever: Failed one or more tests, then passed, but did not try any FPI
7. Nonmaster: Tried but never passed any tests
8. Striving Nonmaster: Nonmaster who did Try any FPI
9. Minimalist Nonmaster: Nonmaster who did NOT Try any FPI
### Table 3

**GA4 Segment Types with Explore Segment Overlap Tool: Achievers and Nonmasters Broken Down by Event Types**

<table>
<thead>
<tr>
<th>Event name</th>
<th>Achievers</th>
<th>Nonmasters</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Active users</td>
<td>Conversions</td>
</tr>
<tr>
<td>All Events</td>
<td>51,646</td>
<td>11,212,169</td>
</tr>
<tr>
<td>page_view</td>
<td>51,646</td>
<td>4,338,184</td>
</tr>
<tr>
<td>Plagiarism_Test</td>
<td>51,646</td>
<td>1,015,190</td>
</tr>
<tr>
<td>Test_Feedback</td>
<td>51,646</td>
<td>1,158,205</td>
</tr>
<tr>
<td>Activation</td>
<td>21,416</td>
<td>281,842</td>
</tr>
<tr>
<td>Demonstration</td>
<td>21,202</td>
<td>248,736</td>
</tr>
<tr>
<td>Application</td>
<td>21,362</td>
<td>2,567,369</td>
</tr>
<tr>
<td>Integration</td>
<td>19,998</td>
<td>165,748</td>
</tr>
<tr>
<td>Mastery_Test</td>
<td>21,488</td>
<td>554,001</td>
</tr>
<tr>
<td>Plagiarism_Patterns</td>
<td>33,601</td>
<td>750,641</td>
</tr>
<tr>
<td>Overall Total FPI¹</td>
<td>4,568,337</td>
<td>88.5</td>
</tr>
<tr>
<td>For those who Tried any FPI²</td>
<td>108.7²</td>
<td>0</td>
</tr>
</tbody>
</table>

¹ Means based on 51,646 Traditional and Minimalist Achievers, as well as 23,307 Striving and Minimalist Nonmasters.

² Means exclude 9,600 Minimalist Achievers and 7,448 Minimalist Nonmasters (See Table 2).
Table 4

Analysis of Patterns in Time (APT): Results from Applying Bayes’ Rules to Convert Joint Probabilities to Conditional Probabilities

<table>
<thead>
<tr>
<th>First Principle of Instruction (FPI)</th>
<th>( p(A &amp; FPI) )</th>
<th>( p(N &amp; FPI) )</th>
<th>( p(A \mid FPI) )</th>
<th>( p(N \mid FPI) )</th>
<th>Odds (A:N)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Activation</td>
<td>0.285</td>
<td>0.081</td>
<td>0.778</td>
<td>0.222</td>
<td>3.513</td>
</tr>
<tr>
<td>Demonstration</td>
<td>0.282</td>
<td>0.078</td>
<td>0.783</td>
<td>0.217</td>
<td>3.618</td>
</tr>
<tr>
<td>Application</td>
<td>0.284</td>
<td>0.075</td>
<td>0.792</td>
<td>0.208</td>
<td>3.818</td>
</tr>
<tr>
<td>Integration</td>
<td>0.266</td>
<td>0.067</td>
<td>0.798</td>
<td>0.202</td>
<td>3.953</td>
</tr>
<tr>
<td>Mastery Test*</td>
<td>0.286</td>
<td>0.078</td>
<td>0.786</td>
<td>0.214</td>
<td>3.670</td>
</tr>
<tr>
<td>Plagiarism Patterns*</td>
<td>0.447</td>
<td>0.180</td>
<td>0.714</td>
<td>0.286</td>
<td>2.492</td>
</tr>
</tbody>
</table>

*Note: The Mastery Test is a further instantiation of the Application principle. Plagiarism Patterns were designed with the Demonstration principle. Achievers (A) failed one or more Certification Tests (CTs), then passed one. Nonmasters (N) failed two or more CTs, never passing one.

Key:

\( p(A \& FPI) \): probability of Achiever and choosing a First Principle of Instruction

\( p(N \& FPI) \): probability of Nonmaster and choosing a First Principle of Instruction

\( p(A \mid FPI) \): probability of Achiever, given that they chose a FPI

\( p(N \mid FPI) \): probability of Nonmaster, given that they chose a FPI

Odds (A:N): odds of being an Achiever vs. a Nonmaster, given that they chose a FPI

GA4 Analysis Reports. Once new events had been created and tracked, as described above, the real power of GA4 could be utilized with its ‘Explore’ tool. We used the ‘Segment Overlap’ analytic technique to do APT of IPTAT usage. Since many readers are likely to be unfamiliar with this approach to data analysis, Frick et al. (2021) have provided video demonstrations which illustrate how to do this with both Universal Analytics and GA4 at https://plagiarism.iu.edu/apt/demo/ .
Results

Doing APT Queries with GA4

Figures 2 to 8 and Tables 1 to 4 provide results from GA4 Segment Overlap analyses for IPTAT users between January 1 and March 25, 2021. As a general strategy, we did retrodictive APT queries. As Frick et al. (2022) explained, retrodictive queries specify an end event and then identify conditions that temporally precede the end event, counting instances which match the pattern specified in the APT query.

In mathematical set theory, two sets are mutually exclusive if their intersection is the empty set; and they are exhaustive if every element can be placed in one of the sets. Thus, every element is a member of one and only one set. Here the elements are learning journeys. APT requires such classifications in order to form mathematical likelihoods. See Frick (1983, pp. 100-104). Moreover, when a given element (learning journey here) is a member of two or more sets, it is contained within the intersection of those sets.

Figure 2 illustrates a subset relationship between two sets, and Figure 8 illustrates an intersection relationship. In GA4 we can define segments, that is, develop criteria which indicate which elements are members of that set. Google calls these ‘segments’, Frick refers to them as ‘phrase segments’ in APT queries (Frick, 1983; Myers & Frick, 2015), mathematicians call them ‘sets’, and computer scientists call them ‘Boolean logic’ for conditional expressions in software engineering. But these are essentially the same basic idea—ways of grouping cases according to criteria for membership in the group, whether or not a case satisfies certain conditions.

Frick et al. (2022) identified three basic groups of students who use IPTAT: Traditionalists, Minimalists, and Dabblers. We noticed in our present study that a sizeable
portion of IPTAT visitors spend relatively little time on the site. IPTAT was designed so that it can be used flexibly. For example, a student can learn from one or more tutorials that were designed with First Principles of Instruction (FPI). They do not need to register in IPTAT to take Certification Tests (CTs), if they choose not to.

Frick et al. (2022) further found that approximately 94% of students who do register for IPTAT indicate that they are there because their instructor or school requires them to pass a CT and provide their Certificate as evidence that they can recognize plagiarism. Frick et al. (2022) further reported that students, on average, make 7 to 8 attempts at unique Certification Tests before they pass. Certificate logs stored at IU, which are records of each Certificate granted (Frick & Dagli, 2016), revealed that once a student passes a CT, they very rarely try and pass further CTs. It is extremely uncommon (estimated as about one in 10,000 Certificates) for a student with the same name and email address who has two or more unique Certificate log entries. Each Certificate has a unique Test ID, and there are trillions of unique Certification Tests.

Therefore, by deduction, when a student has attempted two or more CTs, they must have failed at least one CT before they passed a new one. Moreover, the IPTAT ‘test attempts’ log at IU confirms this pattern. Students typically fail 6 to 7 CTs before finally passing one.

For these reasons, we created a ‘Test Evaluations’ segment in GA4, which is defined by several criteria. To belong to this set, a student must have viewed both an IPTAT webpage for taking a test (‘plagiarismTest*.php’) and received feedback (‘evaluateAnswers*.php’) and there must be more than one pageview of each of these two webpages. Figure 2 shows that the ‘Test Evaluations’ segment is a subset of all IPTAT Active users, consisting of 75,087 learning journeys, and the complement of this subset is those IPTAT users who did not take two or more
CTs and receive feedback \((n = 97,029\) learning journeys). We subsequently chose to focus on the Test Evaluations segment. Note that there are separate test levels for undergraduate and advanced high school students (UG) and graduate level students (GR).

We then identified IPTAT Active Users in the Test Evaluations segment who passed a CT (termed ‘Achievers’, \(n = 51,646\)) and those who did not (‘Nonmasters’, \(n = 23,307\)). These are mutually exclusive and exhaustive subsets of the Test Evaluations set. See Figure 3. Achievers therefore must have failed at least one CT before passing a new one, which is an indication of student learning achievement—at time 1 they were nonmasters, and at a later time 2, they were masters). Nonmasters, on the other hand, by definition have not passed any CT (in this timeframe, although they might have done so after March 25).

IPTAT keeps records at IU of every registered student who passes a CT, so that they and their instructors can later retrieve their unique Certificate. In Google Analytics, there are no tracked pageviews which indicate that a student has failed a test (by IPTAT website design, since this is a privacy issue). However, if a student passes a CT (also not tracked by GA), they can then click a button on the CT feedback page which mails their Certificate to their own email (‘mailCertificate*.php’). GA4 does not know the email address of that student, but can only track when the ‘mailCertificate*.php’ webpage is viewed. IPTAT software engineers have carefully designed the code so that the ‘mailCertificate*.php’ file can only be viewed immediately after passing a CT. They have gone to considerable lengths to engineer proprietary software to prevent any other access to the ‘mailCertificate*.php’ webpage to prevent hackers from generating bogus Certificates (see Frick & Dagli, 2016).

Furthermore, IPTAT software prevents students from trying the same CT again, modifying their answers, and getting feedback on it more than once. They only get one
evaluation of each unique test. They can take multiple CTs, but each is unique—there are literally trillions of unique CTs, again by design. The chance of passing a CT by guessing is one in 59,049. See Frick et al. (2018).

IPTAT software further allows users and their teachers to subsequently validate each unique Certificate. This helps prevent abuse of IPTAT and student cheating. And there are built-in monitoring strategies which trigger alarms (to the developers) when suspicious activity occurs on the IPTAT website. IPTAT software engineers regularly monitor the CT passing rate and have noted that it has consistently remained around 13 to 14 percent since 2016 (Frick et al., 2022).

Finally, Frick et al. (2022) reported that GA pageviews of ‘mailCertificate*.php’ were a small underestimate of the actual number of Certificates issued (about 2 percent less). The primary reasons are either that a student does not click on the button to email their Certificate, or that they are using browsers or browser privacy settings to prevent GA tracking of IPTAT website usage. Nonetheless, the GA reports of ‘mailCertificate*.php’ pageviews are a good indicator of successful student achievement, even if a small underestimate.

Figures 4 through 8 show a further GA4 segment overlap analysis. In addition to Test Evaluations and Achievers segments, a third segment was defined to classify IPTAT users who had tried one or more webpages that were designed with First Principles of Instruction (called ‘Try any FPI’). We did this because Frick et al. (2022) had distinguished Minimalists from Traditionalists.

Minimalists apparently intend to do the least amount possible to pass a test. IPTAT CTs contain embedded links to a webpage (https://plagiarism.iu.edu/hints.html), which provides the rules for classifying kinds of plagiarism and further links to other helpful materials at the bottom.
Except for the ‘Plagiarism_Patterns’ links, these further aids were not designed specifically with FPI in mind, but rather have been added over time to support the Minimalists.

Moreover, each CT feedback page provides tailored results from the test just taken. The results indicate whether that student has passed or not. If not, test feedback is adapted to provide further clues (such as “You answered less than half the questions correctly.” Or “You answered the majority of questions correctly.”); and test feedback contains links to the types of mistakes made. Minimalists do not click on any of those ‘Plagiarism_Patterns’ links, but rather rely on trying more CTs and the ‘hints’ webpage. Frick et al. (2022) have previously confirmed this strategy also by using GA tools for reporting “behavior flows” and by inspection of individual temporal maps.

If students do click on any of the ‘Plagiarism_Patterns’ links, then they are classified as Traditionalists, since the patterns are examples of different kinds of plagiarism—the FPI Demonstration principle was used to design these 19 IPTAT webpages.

Table 4 provides answers to the main research question. It indicates that students are nearly 4 times more likely to pass a Certification Test when they have chosen IPTAT instruction designed with FPI. Tables 1 to 3 provide intermediate data which allowed us to reach this conclusion.

**Segmenting Three Sets of IPTAT Active Users**

Table 2 summarizes results from the three-way Segment Overlap analysis (Test Evaluations x Achievers x Try any FPI). In particular, rows 4 and 7 indicate the main findings. The 51,646 Achievers spent nearly 4.1 million minutes of engagement with the IPTAT website, averaging about 79 minutes each. Achievers also viewed over 5.6 million IPTAT webpages, for an average of 108.5 pageviews during those 79 minutes.
On the other hand, the 23,307 Nonmasters spent an average of 66 minutes on IPTAT, viewing 79.3 webpages. Thus, Achievers viewed on average 29.2 more IPTAT webpages than did Nonmasters and spent about 13 more minutes engaged with IPTAT.

In Table 2, rows 5 and 6, Achievers were further broken down into segments for those who had tried any FPI-designed webpages (Traditionalist Achievers, \( n = 42,046 \)) and those who had not (Minimalist Achievers, \( n = 9,600 \)). Traditionalist Achievers spent about 90.4 minutes on average, viewing 128.5 IPTAT webpages; whereas Minimalist Achievers spent about 28.9 minutes and viewed 21.2 webpages on average. Traditionalist Achievers spent over three times as much time than did Minimalists who also passed CTs.

Finally, Table 2 row 9 indicates that 7,448 Minimalist Nonmasters spent about 25 minutes on average, viewing about 19 IPTAT webpages (none of which were designed with FPI). Alternatively, row 8 indicates that 16,032 Nonmasters had tried FPI-designed webpages but still had not passed a CT. We call this group Striving Nonmasters. These students spent about 84.5 minutes on average, viewing about 106.3 webpages overall, while not passing any Certification Tests. However, when we focus specifically on FPI-designed webpages, this group averaged 75.9 pageviews, as indicated in Table 3, whereas the Traditionalist Achievers viewed an average of 108.7 FPI-designed webpages. The Traditionalist Achievers had viewed on average 32.6 more FPI-designed webpages when compared with the Striving Nonmasters.

The Striving Nonmasters appear to be trying to learn, but they have not achieved mastery in this study’s timeframe. Some of these students may later pass a Certification Test. Indeed, we have subsequently determined from IU MySQL records that there were 5,976 students who registered on or before March 25 but passed a CT after March 25 (as of December 16, 2021, 11:40 a.m. EST).
Overall, Table 2 indicates that about 80 percent of the Achievers (42,046) had tried one or more IPTAT webpages designed with First Principles of Instruction. The other 20 percent of Achievers had utilized the test ‘hints’ page and test feedback as their primary strategy until they eventually passed a CT. These Minimalist Achievers avoided the IPTAT tutorials and plagiarism patterns webpages, spending about one-third as much time overall when compared to the Traditionalist Achievers.

The 23,307 Nonmasters constituted about 31 percent of the IPTAT users who took more than one Certification Test and received test feedback, compared with the 69 percent who were Achievers. Frick et al. (2022) reported that from 2016 through 2020, about 81 percent of those who registered to take CTs did pass one during that time frame, according to IU MySQL records on Certificates granted. As discussed above, nearly 6,000 of these Nonmasters eventually passed a CT after March 25, 2021, and thus those 25.6 percent became Achievers after the timeframe of our study, which is based on GA4 session records.

FPI-Designed IPTAT Web Pageviews by Achievers versus Nonmasters

Table 3 provides further breakdowns by GA4 event names for Achievers and Nonmasters, and, in particular, for the different First Principles of Instruction. Noteworthy is that Achievers and Nonmasters appear to have tried about the same number of Certification Tests (‘Plagiarism_Test’) and received feedback on results (‘Evaluate_Answers’). Means of pageviews are in the 20 to 23 range.

Further noteworthy are the numbers of Active Users who experienced various First Principles in the Achievers segment, roughly between 20K and 21K, while for the Nonmasters segment there were between 5K and 6K Active Users. Achievers viewed webpages designed with the Application principle over 2.5 million times, for an average of 49.7 per user. On the
other hand, Nonmasters viewed 618K Application webpages, for an average of 26.5 per user. Achievers viewed, on average, almost twice as many Application webpages (practice questions with explanatory feedback) as did Nonmasters. Achievers likewise viewed nearly twice as many pages on average for the remaining First Principles (Activation, Demonstration, Integration, and Mastery Test). However, those averages were much smaller, though ratios were similar (e.g., for Activation pageviews per Achiever vs. Nonmaster: $5.5/2.9 = 1.9$; for Demonstration pageviews per Achiever vs. Nonmaster: $4.8/2.7 = 1.8$). See Table 3.

Achievers viewed more than 4.5 million FPI-designed webpages overall, for an average of 88.5 pageviews. In contrast, Nonmasters viewed just over 1.2 million FPI-designed webpages, for an average of 52.2 per Active User. This is a difference of 36.3 more FPI-designed pages that were viewed by Achievers when compared with Nonmasters.

Note further that the numbers of Achievers and Nonmasters were fairly similar for nearly all First Principles, with the exception of ‘Plagiarism_Patterns’. The categories of First Principles are not mutually exclusive, so the same users could have experienced more than one of the First Principles. As can be seen in Table 2, there were 42,046 Active Users who experienced one or more First Principles. In Table 3 there were about 21K users who experienced each type of First Principle (Activation through Mastery_Test) for a total of about 105.5K (and $105.5K/42.0K = 2.5$). Thus, we can infer that each Active User who was a Traditionalist Achiever viewed 2.5 of the 5 types of FPI-designed webpages. For Striving Nonmasters, there were about 28.5K total Active Users, divided by about 16.0K, or about 1.8 types of FPI-designed webpages. Therefore, the Traditionalist Achievers experienced about one more type of First Principle than did Traditionalist Nonmasters. The Minimalist Achievers and Minimalist Nonmasters experienced zero types of FPI-designed webpages.
Likelihoods of Passing a Certification Test when First Principles of Instruction were Experienced

Table 4 provides results derived from Table 3 APT frequency counts. Bayes’ rules were used to convert to relative frequencies. For example, for the Activation Principle there were 21,416 Active Users who were Achievers out of the 75,087 who took more than one Certification Test and received feedback on results. Achievers, by definition, failed one or more CTs before passing. Thus, the likelihood of the joint occurrence of failing then passing a CT and experiencing the Activation principle is 21,416/75,087 or 0.285. Similarly, the likelihood of the joint occurrence of not passing a CT and experiencing the Activation principle is 6,097/75,087 or 0.081. When Bayes’ rule is applied, these joint probabilities are converted to conditional probabilities. Thus, we can conclude that the likelihood of failing then passing a Certification Test, given that the Activation principle is experienced one or more times is 0.778 (0.285/(0.285 + 0.081)). The likelihood of failure is 0.222 when the Activation principle is experienced one or more times (1 − p(A | FPI=Activation)). This can be interpreted: the odds of passing versus failing an IPTAT Certification Test are about 4 times greater when instruction designed with the Activation principle is selected.

Similar odds were observed for Demonstration, Application, Integration and Mastery Tests. The overall odds for passing versus failing a CT, given that any of the FPI-designed tutorials were experienced, were about 3.7 to 1. The odds of passing versus failing a CT when Plagiarism Patterns were experienced were somewhat less, about 2.5 to 1.

We further investigated combinations of FPI types. We found that 29,531 Active Users experienced Activation and Demonstration and Application together, which was the largest number of active users for various combinations of FPI types. In a further Segment Overlap
analysis, we found that there were 19,104 of those 29,531 Active Users who were Achievers, in contrast to 4,992 Nonmasters (who took 2 or more CTs). This is an odds ratio of 19,104/4,992 or 3.8 to 1. That is, when students experience Activation, Demonstration, and Application (at least once each), they are about 3.8 times more likely to pass a CT. Note that the remaining 5,458 Active Users did not take 2 or more Certification Tests.

We further know that about 20 percent of the Achievers who were Minimalists did not choose any of the FPI-designed webpages; and they were able to pass tests by trying multiple times, receiving feedback on test results and by using test hints. The remaining 80 percent of Achievers who were Traditionalists did appear to benefit from their choice of webpages designed with First Principles of Instruction. See Table 2.

Frick et al. (2022) reported that, historically, students occasionally complain about the difficulty of the CTs and the fact that they are not provided with correct answers to questions missed on the test. This is again by design, a necessary compromise to accommodate instructor concerns about widespread cheating (as occurred in 2012, but was addressed in 2013 by creating trillions of unique CTs: https://plagiarism.iu.edu/recentChanges.html). IPTAT designers have since developed a frequently asked questions (FAQ) page to discuss this issue, especially https://plagiarism.iu.edu/faq.html#faq2 and https://plagiarism.iu.edu/faq.html#faq16.

Frick et al. (2022) reported that from both anecdotal evidence (from students who complain how hard the tests are) and by examining behavior flows in GA, the initial Minimalist strategy of trying and repeatedly failing numerous tests often proves unsuccessful. Those frustrated students eventually conclude that they need to do some of the tutorials with practice tests. If, when they return to IPTAT, they are using a different device with a different IP address and potentially a different Web browser, GA may identify them as a new Active User. In other
words, an initial Minimalist Nonmaster could also turn into a Traditionalist Achiever. These could appear to be different learning journeys, based on this limitation of Google Analytics in how unique client-IDs are generated for Active Users.

Despite this GA limitation, the fact remains that the odds are nearly 4 times greater for passing a CT if one or more tutorial webpages designed with First Principles of Instruction are chosen by students during their learning journeys.

We also note from our results that those Minimalist Achievers are literally “quick learners” who apparently require minimal instruction. As Table 3 indicates, they spend an average of 28.9 minutes engaged with the IPTAT website, compared with about three times as much time spent by Traditionalist Achievers. These Minimalist Achievers apparently do not need much guidance in passing a CT. Test hints and test feedback appear to be enough guidance for them to pass a CT, while the other 80 percent of Achievers benefit from FPI-designed instruction.

Finally, we note that about 85 percent of students who use IPTAT are undergraduate or graduate students in higher education. Presumably they are more capable learners with higher levels of intelligence (i.e., greater ability to learn). Those Minimalist Achievers may be very high ability learners, which a further research study might be able to verify. They also may have a better understanding of plagiarism before they experience any of IPTAT, when compared to Traditionlist Achievers and Nonmasters.

**Discussion**

**Generalizability of Findings**

Results from the present study are similar to those from an earlier two-year Big Study that Frick et al. (2022) conducted on IPTAT using a prior version of Google Analytics, known as
Universal Analytics (UA). On the other hand, we used GA4, a relatively new version of GA that became available in mid-October 2020.

In the Frick et al. (2022) study, they reported results on more than 936,000 learning journeys. According to UA tracking, those students were located in 213 countries and territories worldwide. Those 390,000+ users who registered for IPTAT reported ages mostly between 14 and 44 years old. In the present study, GA4 tracked more than 172,000 IPTAT learning journeys (Active Users) for nearly three months, and we focused on Nonmasters and Achievers. According to GA4 tracking, these users were located in 186 different countries and territories worldwide. In both studies, when students registered for IPTAT, the large majority (94 percent) reported that they were doing so because it was a requirement by their teacher or school; and from their reported ages, we can infer that they were mostly college and advanced high school students (adults).

Frick et al. (2022) reported that successful students were between three and four times as likely to choose unique webpages designed with First Principles, when compared with those who did not pass Certification Tests. On the other hand, with GA4 we were able to determine the likelihood of passing a CT when FPI-designed webpages were selected by students. We could take advantage of GA4’s better tracking capabilities and more powerful analysis tools (especially Segment Overlap). Specifically, results in Table 3 could be reported with GA4, e.g., that 21,416 Active Users experienced the Activation principle at least once, for a total of 281,842 pageviews (marked as GA4 Conversions). This kind of specific connectivity of Active Users with Conversions was not possible in Universal Analytics, which is what Frick et al. (2022) had utilized. They were able to refer to unique pageviews of FPI-designed webpages by an Audience Segment (the overall number of Active Users in that segment). With GA4 we were able to
obtain a *per Active User* metric for each First Principle in each Audience Segment. This then allowed us to estimate the Bayesian likelihoods of passing a CT (i.e., student achievement) when each First Principle of Instruction was chosen by IPTAT students (see Tables 3 and 4). Frick et al. (2022) could not make such estimates. Instead of determining odds of choosing webpages as they did, we were able to calculate odds of successful learning journeys when some FPI-designed webpages were selected.

In the Frick et al. (2022) study, they segmented by Certificants and Non-Certificants. Their Non-Certificant segment included Dabblers, who were not excluded when using Universal Analytics. With GA4, we were able to clearly separate the Dabblers from the rest. Dabblers spent very little time in our study (mean of 3.4 minutes per user, and 8.5 pageviews), yet constituted over half the Active Users (97,029 out of 172,417 = 56.2 percent). With GA4, we were able to focus on students who were attempting to pass Certification Tests, by excluding the Dabblers. Frick et al. (2022) may be able to re-analyze their data in a similar way with Universal Analytics, now that we have shown the value of doing so with GA4.

A further difference between the present study and Frick et al. (2022) is the average amount of engagement time per learning journey. In our study, we found that Achievers averaged about 79 minutes per journey, whereas Frick et al. reported a mean of about 98 minutes for successful students. We found in our study that Nonmasters spent an average of 66 minutes actively engaged in using IPTAT.

We note further that GA4 tracks user engagement somewhat more precisely—referring to *engaged sessions*. GA4 can apparently discriminate whether the app being used is the user’s focus. When a user switches away from the app being tracked, GA4 does not consider this to be active user engagement. For example, during IPTAT, a student might switch away to read and
respond to their text messages or e-mail for several minutes, and then resume interacting with IPTAT. In UA this would not be specifically tracked, unless the user had switched away for more than 30 minutes, in which case the IPTAT session would be terminated (as it also would be in GA4). However, in GA4 engaged time will apparently only accrue when the app being tracked is the user’s focus. We believe that this is the most likely explanation of the differences between our results and those of Frick et al. (2022). Thus, we conclude here that the GA4 results are better indicators than UA when student engagement time is being considered.

The main difference between our study and Frick et al. (2022) is that we excluded those who did not try at least two Certification Tests (CT) and receive feedback on results (referred to as Dabblers). Note further that Dabbler learning journeys constituted about 56 percent of the total learning journeys. The remaining 44 percent of journeys were much longer in duration (between 66 and 79 minutes of engagement, with an average of 79 to 109 pageviews per Active User). See Table 1.

Very importantly, Achievers in our study demonstrably showed evidence of an increase in learning, whereas Nonmasters had not passed any CTs during the interval in which they were observed (who attempted two or more CTs between Jan. 1 and Mar. 25, 2021).

**Temporal Retrodiction from APT, Not Causal Inferences**

Frick (1990) noted that causal inferences are not warranted from APT results unless additional factors are considered. He provided the example of dawn and sunrise. Dawn is a good predictor of sunrise, but dawn does not necessarily cause sunrise. Scientific theory that involves Newton’s laws and optical refraction of light is preferable for explaining cause and effect of dawn and sunrise. Nonetheless, sunrise is highly predictable following dawn—for
those of us not living near the North or South Poles. And we can make decisions based on predictability of events, even if we cannot provide causal explanations.

Thus, we should not conclude that First Principles of Instruction cause student learning success (passing a Certification Test). But from APT results, we can conclude that, when users have selected and utilized at least some instructional components designed with First Principles, they are nearly four times more likely to pass than fail a Certification Test.

APT queries can be predictive or retrodictive, as explained by Frick et al. (2022). Based on present conditions, prediction is forecasting what is likely to happen in the future. On the other hand, retrodiction is looking backward in time. Given that some event has been observed, what events happened at an earlier time? This is what we did in the present study. We observed when students passed a Certification Test during their learning journeys and compared them to others who had not passed in their learning journeys. And we further selected those who had received feedback on a Certification Test at least twice. Then we looked backwards in time and counted how often each of these two groups chose parts of IPTAT designed with FPI. Myers and Frick (2015) likewise did this in their study of the Diffusion Simulation Game. They classified how well each student had done at the end of a game, categorizing them according to how many adopters each player had achieved. Then they observed prior game strategies players utilized. This also was a retrodictive approach to APT.

In short, APT is a descriptive-correlational approach to empirical research, not unlike what epidemiologists do when attempting to predict whether someone is likely to get cancer. What events precede cancer? For decades, epidemiologists and medical scientists knew that people who smoked cigarettes heavily earlier in their lives were between five and ten times more likely to contract lung cancer later in their lives than were nonsmokers. However, proving the
causal relationship was more challenging until the role of carcinogens was identified as a significant factor, when researchers had a better understanding of biochemistry, pathology, and cell biology. Nonetheless, people were advised to refrain from smoking based on the temporal relationship.

Causal inferences are not warranted by correlation alone, as research methodologists and statisticians have reminded us for decades (e.g., Kirk, 1999). Nonetheless, there are ways to make causal inferences from unintended experiments. In the New York Times, Jeanna Smialek reported on three recent winners of the Nobel Memorial Prize in Economic Sciences. Smialek quoted Peter Fredriksson, who chaired the Nobel Prize Committee: “Uncovering causal relationships is a major challenge… Sometimes, nature, or policy changes, provide situations that resemble randomized experiments. This year’s laureates have shown that such natural experiments help answer important questions for society” (Smialek, 2021).

While we do not make causal claims in the present study, we do note that what has happened with IPTAT over the past 19 years could be conceived as a natural experiment as do researchers in economics. We emphasize here that APT is a research methodology which is distinctly different from the linear models approach (LMA) often used by economists.

**APT and Praxiological Theory**

Analysis of Patterns in Time does not identify patterns all by itself. Researchers must specify APT queries, and then APT accordingly segments temporal maps, finds matches, and counts occurrences of those events. While examining individual temporal maps may provide leads on what patterns to tell APT to look for, theory should be driving the process. Kurt Lewin has been often quoted for saying, “Nothing is as practical as a good theory” (Greenwood & Levin, 1998, p. 19).
The present study was guided by instructional theory—in particular, First Principles of Instruction. And design of instruction was driven by a practical need: IPTAT was designed with the goal of helping students learn to recognize basic kinds of plagiarism from non-plagiarism. Designers wanted IPTAT to be effective, that is, achieve its goal, and they wanted it to work via the Web so it would be easily accessible by students in the Instructional Systems Technology program at Indiana University. When IPTAT was designed in 2002, instructional theory about how to teach concepts was applied. In 2015, when IPTAT was redesigned, the design team specifically chose First Principles of Instruction for theoretical guidance. And most importantly, FPI theory not only influenced how they structured the IPTAT website and named webpages in 2015, it also guided which patterns we specified in APT to be counted in the present study.

The results provided in Tables 1-4 did not “emerge from the data,” nor did artificial intelligence algorithms discover those patterns. The segment event types listed in Tables 3 and 4 are identified by names of First Principles. The additional mastery tests and Certification Tests were designed to assess how well students had learned to recognize plagiarism. The tests themselves and feedback on test results are further instantiations of the FPI Application principle. The plagiarism patterns to which test feedback adaptively links are further instances of the FPI Demonstration principle.

In short, theory not only guided the instructional design of IPTAT, that instructional theory also guided the search for patterns of instructional effectiveness. APT was the particular research methodology that guided how to find those patterns, and APT in turn was developed retroductively from general systems theory, information theory, and set theory and probability theory from mathematics (Frick, 1983, 1990). The patterns themselves are qualitative—the event patterns are not numbers, rather they are named. The results of APT are quantitative—
numerical counts of patterns of event occurrences in temporal maps. Those counts are used to form proportions (or likelihoods) and likelihood ratios (odds).

Numerous researchers have noted the need to leverage the power of learning analytics for evaluating instructional designs (Gašević et al., 2015; Ifenthaler, 2017; Klein & Hess, 2019; Mangaroska & Giannakos, 2019; Phillips & Ozogul, 2020). By explicitly designing the IPTAT using FPI and showing that successful learners experienced more instances of instruction based on FPI than unsuccessful learners, we have added support to Merrill’s hypothesis regarding the effectiveness of FPI while also demonstrating how learning analytics (in the form of APT) can be used to test instructional design theory. Recent reviews of learning analytics research have noted the need to integrate education theories and learning analytics (Phillips & Ozogul, 2020; Romero & Ventura, 2020; Wong et al., 2019), and the present study illustrates a pathway toward “a synergistic relationship between instructional design and learning analytics” (Ifenthaler, 2017, p. 202).

**Conclusion**

If generalizability of empirical research results is considered with respect to repeatability of findings and applicability to a wide range of student learners, then results from the present study and the Frick et al. (2022) study can be considered as highly generalizable. When taken together, based on more than two years of IPTAT usage, these APT results apply to a worldwide audience of adult learners who can read and comprehend English and who have computer technology to access the World Wide Web. These observed learning patterns have been highly consistent since 2016, as Frick et al. (2022) have noted.
Further Comparative Studies Using APT

As Merrill noted about the Frick et al. (2022) study, “While [it] does not provide a direct comparison of instruction based on these principles with instruction not based on these principles, it nevertheless demonstrates that instruction based on First Principles makes it far more likely that a student will pass the test” (Foreword in Frick et al., 2022, p. x). The present study is similar to Frick et al. (2022) in this regard. Both are comparing how students use the instructional materials in various ways that are associated with passing a Certification Test.

These studies are not comparing instruction designed with First Principles to other instruction based on different principles. They do not allow claims such as First Principles are more (or less) effective than other principles used in design of instruction. Some readers may view this as a limitation of these studies. The designers’ goal in developing IPTAT was not to do a comparative study; rather those designers wanted IPTAT to be effective, efficient, and highly flexible. By design, students are free to choose whatever parts of the instruction they want or need to pursue when learning. In other words, IPTAT is flexible and adaptable.

Nonetheless, it should be patent that comparative studies could be done using APT methods. For example, if successful students in approach X were 5 times more likely to choose instruction based on X principles, when compared to successful students in approach Y who were 10 times more likely to choose instruction based on Y principles, then we could claim that instructional components designed with approach Y are twice as likely to be chosen by successful students as that with approach X. Or if approach X resulted in a likelihood of student success of 0.80, whereas approach Y resulted in a likelihood of success of 0.85, we could say that while approach Y is 5 percent more effective than approach X, the latter is much more efficient since students spent about half as much time learning, and yet were nearly as successful overall.
One of the ETR&D reviewers suggested that we analyze *sequences of pageviews*, as a further way to characterize IPTAT usage. GA does provide an analytic tool to investigate unique sequential paths through a website, referred to as ‘behavior flows’ in UA and ‘path explorations’ in GA4. When we have investigated paths with more than a few vertices in the digraph, it quickly becomes apparent that the “Other” group becomes increasingly large at each additional vertex (path step). This should not be surprising, since IPTAT has high degree of flexibility with respect to possible navigation paths. For example, users can click on the Sitemap link on nearly every IPTAT webpage, which contains direct links to over 80 other pages. Or they can click on nine other major sections in the navigation sidebar, and then choose from further submenus within each section. Thus, most webpages are accessible from any other IPTAT webpage by selecting the sidebar Sitemap and one of its links, or by selecting one of the other sidebar menus, and two further links. If IPTAT instruction is followed in the designed sequence, each webpage is directly accessible via the “Next Page” link. Any other given webpage is usually two to three clicks or taps away. The number of unique permutations of sequential paths through IPTAT is extremely large.

In the present study, we have focused on a relatively small subset of temporal patterns in Tables 2, 3, and 4 which are relevant to student choice of instruction designed with First Principles. We have captured the prominent global patterns of usage, and have named them as Dabblers, Traditionalist Achievers, Minimalist Achievers, Striving Nonmasters, and Minimalist Nonmasters.

**Final Remarks**

The primary goal of the Indiana University Plagiarism Tutorials and Tests has been to help students learn to recognize plagiarism. IPTAT has been available as a MOOC since 2002 at
no charge to users. More than 1,113,724 students worldwide have successfully registered to take Certification Tests in IPTAT between Jan. 2, 2016 and noon on December 16, 2021. Over 900,004 have passed a CT in that same timeframe. In the present study, students who viewed IPTAT tutorials designed with FPI were four times more likely to become Achievers, when compared with Nonmasters who had tried but failed to pass Certification Tests. These results are an indicator of the effectiveness of First Principles of Instruction with respect to their instrumental value.

While we do not conclude that FPI causes student learning achievement, we have observed repeatedly over a period of nearly six years that students who succeed in their learning are much more likely to choose parts of IPTAT designed with FPI. We do not believe that this is a coincidence or a random occurrence, since the likelihood ratios have remained relatively stable, and because there has been a very large number of students from all over the world who have achieved mastery in recognizing word-for-word and paraphrasing plagiarism from non-plagiarism.

Compliance with Ethical Standards

- Disclosure of potential conflicts of interest. The authors have no conflicts of interest to declare that are relevant to the content of this article. This study was not funded. The authors did not receive support from any organization for conducting this study.
- Research involving human participants and/or animals. This study has been approved and granted exemption for human subjects research by the Indiana University Institutional Review Board, Protocol No. 1304011238.
- Informed consent. No informed consent was required for this study by the Indiana University Institutional Review Board, Protocol No. 1304011238. The Privacy Policy for
the Indiana University Plagiarism Tutorials and Tests is stated at

https://plagiarism.iu.edu/privacy.html. In compliance with the Privacy Policy, we share only aggregate, non-personally identifiable information about participants in this study.

References


[https://support.google.com/analytics/answer/10089681?hl=en](https://support.google.com/analytics/answer/10089681?hl=en)


Merrill, M. D. (2020). *M. David Merrill’s First Principles of Instruction*. Association for Educational Communications and Technology.


Steiner, E. (2009). Ethical theory (Lecture notes).

https://tedfrick.sitehost.iu.edu/steiner/Ethical%20Theory.pdf

