***Приложение А.*** Описание процедуры статистической обработки данных.
***Appendix A.*** Description of the statistical data processing procedure.

## Описание процедуры статистической обработки данных
## Description of the statistical data processing procedure

С целью выявления латентных профилей учебной активности студентов на основе «цифровых следов», оставленных в ЭОС, использовался метод анализа латентных профилей (Latent profile analysis, LPA). Данный метод является разновидностью моделирования латентных классов (latent class modeling), позволяющей на основании набора эмпирических данных выделять латентные классы или профили популяции, обладающие сходным набором характеристик. В отличие от кластерного анализа LPA основан на вероятностном подходе, поэтому он позволяет лучше понять структуру сложных популяций. В настоящее время LPA широко используется в психологии, медицине, образовании, маркетинге и т.д.

Обработка полученных данных проводилась с помощью языка программирования R (версия 4.5.1) в среде RStudio, с применением следующих библиотек и статистических пакетов: «missForest» – для восполнения пропущенных данных с использованием метода, основанного на случайных лесах; «tidyLPA» – для анализа латентных профилей (LPA). Для определения пригодности построенных моделей использовались следующие показатели Log-Likelihood (LogLik) – мера соответствия модели данным; информационные критерии: Akaike Information Criterion – AIC, Bayesian Information Criterion – BIC, Энтропия (Entropy) – оценка качества классификации (чем ближе к 1, тем лучше), BLRT (Bootstrapped Likelihood Ratio Test), который проверяет, улучшает ли добавление нового класса модель (если $p < 0.05$, то добавление нового класса статистически значимо улучшает модель).

Для выявления значимых предикторов латентных профилей использовалась функция «stepclass», оценивающая точность (correctness rate), с которой отобранный ею набор предикторов предсказывает принадлежность респондента к латентным классам, выявленным с помощью LPA. Для интерпретации величины точности классификационной модели (точности отнесения респондента в определённый класс, на основании отобранного набора предикторов) использовался комплексный подход, основанный на трёх критериях: (1) сравнение с базовым уровнем (baseline accuracy) – точность модели сравнивалась с точностью наивного прогноза, в качестве которого принималось отнесение всех наблюдений к наиболее многочисленному классу (maximal class baseline). Accuracy модели ниже 55% свидетельствует о том, что её прогноз ненамного превышает уровень простого угадывания, что ставит под сомнение её практическую полезность; (2) сложившаяся научная конвенция: в литературе по машинному обучению точность в диапазоне 60-70% рассматривается как минимальный порог, начиная с которого модель может считаться имеющей ограниченную практическую применимость; (3) абсолютное значение метрики: точность ниже 55% означает, что модель ошибается более чем в половине случаев, а для задач практического прогнозирования подобный уровень ошибки является неприемлемо высоким.

После анализа латентных профилей нами был применен линейный дискриминантный анализ (Linear Discriminant Analysis, LDA) для выявления ключевых переменных, которые наилучшим образом разделяют уже выделенные классы. Результаты анализа представляются в виде линейных дискриминантных функций, каждая из которых объясняет свой процент различий между классами. Переменные, образующие линейные дискриминантные функции (далее LD1 или LD2) могут быть с положительным или отрицательным знаком, который означает полюс, полученной функции или направление связи переменной с дискриминантной осью (положительный знак указывает на положительный полюс оси). Величина коэффициента показывает силу этой связи. А сами

оси имеют не только математический смысл, но и содержательный (интерпретируемый), извлекающийся путем анализа знаков и величины коэффициентов. Для интерпретации полученных результатов мы опирались на следующие показатели: (1) доля следов (proportion of trace) — показатель, который отвечает на вопрос: «Насколько важна каждая дискриминантная функция для различения всех групп?», где число отражает процент объясненной дисперсии; (2) нагрузки (loadings) — это корреляции предикторов с дискриминантными осями, для определения «практически значимой» нагрузки нами используется порог |0,3| – |0,4|, так как при проведении LDA принято использовать те же значения, что и при факторном анализе (Comrey, Lee, 2013); (3) стандартизованные коэффициенты LDA – несмотря на то, что в научной литературе нет общепринятого порога с целью выявления наиболее весомых предикторов нами рассматривались значения с абсолютной величиной выше |0,10|.

Проверка статистической значимости выявленных с помощью функции «stepclass» (на основе LDA) различий между классами по выявленному на шаге 3 набору информативных предикторов с помощью многомерного дисперсионного анализа (функция «Wilks.test»).

**Для цитирования:** Шляпников, В.Н., Шестова, М.А. (2026). Личностные предикторы стратегий саморегуляции самостоятельной работы студентов при дистанционном обучении. *Экспериментальная психология, 19*(2)*,* 138—154. https://doi.org/10.17759/exppsy.2026190209

**For citation:** Shyapnikov, V.N., Shestova, M.A. (2026). Personality predictors of self-regulation strategies for independent work of students in distance learning. *Experimental Psychology (Russia), 19*(2)*,* 138—154. (In Russ.). https://doi.org/10.17759/exppsy.2026190209