بناء نظام تقييم ذكي بتوظيف خوارزميات تعلم الآلة لتشخيص مهارات التفكير العلمي لدى تلاميذ المرحلة الابتدائية

Developing an Intelligent Assessment System Using Machine-Learning Algorithms to Diagnose Scientific Thinking Skills among Primary School Pupils

د. / أمل التلوانى عبد العظيم

مدرس علم النفس التربوي – كلية التربية جامعة دمنهور

نائب مدير مركز ضمان الجودة – جامعة دمنهور

بناء نظام تقييم ذكي بتوظيف خوارزميات تعلم الآلة لتشخيص مهارات التفكير العلمي لدى تلاميذ المرحلة الابتدائية

د./ أمل التلواني عبد العظيم
مدرس علم النفس التربوي – كلية التربية جامعة دمنهور
نائب مدير مركز ضمان الجودة – جامعة دمنهور

المستخلص:

يهدف هذا البحث إلى بناء نظام تقويم ذكي لتشخيص مهارات التفكير العلمي لدى تلاميذ الصف السادس الابتدائي عبر توظيف خوارزميات تعلم الآلة، حيث طوّرت بطارية متكاملة للأدوات تضم اختبارًا تحصيليًا مقنّنًا، ومهامًا أدائية عملية، وبطاقة ملاحظة سلوكية تغطي ثماني مهارات أساسية (الملاحظة، صياغة الفرضيات، التصميم التجريبي، التنبؤ، تفسير البيانات، الاستنتاج، ضبط المتغيرات، والتواصل العلمي). شملت العينة الأساسية 400 تلميذ/ة من محافظة البحيرة (مصر)، وأظهرت الأدوات صدق محتوى مرتفعًا وثباتًا جيدًا، إضافة إلى اتساقات بنائية منطقية عبر مصفوفة الارتباط، بُني نموذج تصنيف باستخدام خوارزمية الغابة العشوائية لتحديد مستويات التفكير العلمي (ضعيف/متوسط/ممتاز)، حقق النموذج دقة كلية بلغت 87.5%، مع ارتباط موجب قوي بين الدرجة المركّبة والتصنيف ($r\approx0.88$) ، وكشف تحليل التباين (ANOVA)على فروق كبيرة الدلالة وحجوم أثر مرتفعة عبر جميع المهارات. كشفت «أهمية السمات» أن صياغة الفرضيات ثم الملاحظة العملية هما الأكثر إسهامًا في التنبؤ بالمستوى، تليهما تفسير البيانات والتنبؤ والاستنتاج، بينما جاء التواصل العلمي أقل وزنًا نسبيًا. وتشير النتائج إلى جدوى التقويم الذكي في تقديم تشخيص لحظي دقيق يدعم التدخل العلاجي المبكر، والتمهيد للتخصص التعليمي فيما بعد، ويزوّد المعلم بمؤشرات قابلة للاستخدام في تصنيف التلاميذ داخل الفصل الدراسي، كما تبرز توجهات تطبيقية لتضمين مخرجات النمذجة في تطوير المناهج وبرامج تنمية المعلمين.

الكلمات المفتاحية:
التقييم الذكي، تعلم الآلة، مهارات التفكير العلمي

الترقيم الدولي الموحد للطباعة ISSN: 3009-612X
الترقيم الدولي الموحد الالكتروني E. ISSN: 3009-6146
المجلة معرفة على قاعدة المجلات العلمية فى بنك المعرفة المصرى
وقاعدة بيانات Edu Search دار المنظومة

Developing an Intelligent Assessment System Using Machine-Learning Algorithms to Diagnose Scientific Thinking Skills among Primary School Pupils

Abstract:

This study develops an intelligent assessment system to diagnose scientific thinking skills among sixth-grade pupils by leveraging machine-learning techniques. A comprehensive measurement battery was designed comprising a standardized achievement test, performance-based tasks, and an observational checklist covering eight core skills (observation, hypothesis formulation, experimental design, prediction, data interpretation, inference, control of variables, and scientific communication). The main sample included 400 pupils from Al-Beheira (Egypt), with a pilot (n=46) for psychometric verification. The instruments demonstrated strong content validity and high reliability (overall $\alpha \approx 0.91$), alongside coherent inter-skill correlations supporting a composite score. A Random Forest classifier was trained to assign pupils to three proficiency levels (Weak/Average/Excellent). The model achieved 87.5% overall accuracy; the composite score correlated strongly and positively with the final classification ($r \approx 0.88$); and one-way ANOVAs revealed statistically significant differences with large effect sizes across all skills. Feature-importance analysis indicated that hypothesis formulation and practical observation were the most influential predictors, followed by data interpretation, prediction, and inference, with scientific communication contributing least. Findings substantiate the practical value of intelligent assessment for timely, fine-grained diagnosis that enables early remedial actions and personalized learning pathways. The study further outlines actionable implications for classroom dashboards, curriculum refinement, and teacher professional development informed by model outputs.

Keywords:

Intelligent assessment; Machine learning; scientific thinking skills

ISSN: 3009-612X الترقيم الدولي الموحد للطباعة
E. ISSN: 3009-6146 الترقيم الدولي الموحد الالكتروني
المجلة معرفة على قاعدة المجلات العلمية فى بنك المعرفة المصرى
وقاعدة بيانات Edu Search دار المنظومة

مقدمة:

يمثل التفكير العلمي أحد الكفايات الأساسية في تعليم العلوم، وقد شهدت مهارات التفكير العلمي تطورًا ملحوظًا في العقود الأخيرة على مستوى المفهوم والممارسة التربوية، وفي السابق ركزت مناهج العلوم التقليدية على حفظ المعارف والمعلومات، مع إهمال نسبي للمهارات العقلية العليا كالاستقصاء والاستدلال، لكن مع تصاعد الدعوات لإعداد أجيال قادرة على حل المشكلات واتخاذ قرارات مبنية على الدليل، اكتسب تعليم مهارات التفكير العلمي زخمًا متزايدًا.

عرف Zimmerman (2007) التفكير العلمي بصورة واسعة بأنه يشمل المهارات المتضمنة في عمليات الاستقصاء والتجريب وتقييم الأدلة والاستدلال، وذلك بهدف بناء الفهم العلمي وتعديل المفاهيم في ضوء الأدلة، ويشير هذا التعريف إلى أن جوهر التفكير العلمي يكمن في مهارات التفكير والسببية الداعمة لتكوين المفاهيم العلمية وتطويرها، وقد أكدت الأدبيات التربوية منذ تسعينيات القرن الماضي على دمج أنشطة الاستقصاء العلمي في جميع المراحل التعليمية، بما في ذلك المرحلة الابتدائية، لضمان تنمية هذه المهارات لدى المتعلمين منذ سن مبكرة .

كما أظهرت أبحاث علم النفس التربوي وعلم النفس النمو اهتمامًا متزايدًا بفهم كيفية تطور تفكير الأطفال العلمي عبر المراحل العمرية المختلفة، وإمكانية تعزيز هذا التطور من خلال بيئات تعلم ثرية بالتجريب والنقاش العلمي (Zimmerman, 2007). وبحلول القرن الحادي والعشرين، أصبح التفكير العلمي جزءًا أصيلًا من مهارات القرن الحادي والعشرين التي تسعى النظم التعليمية إلى تنميتها مثل التفكير الناقد وحل المشكلات، حيث يُنظر إلى التفكير العلمي على أنه مهارة حياتية تُمكّن الأفراد من فهم العالم واتخاذ قرارات مستنيرة مبنية على الأدلة (Kuhn, 2010).

وعلى مستوى المرحلة الابتدائية في جمهورية مصر العربية، صدرت مناهج علوم محدثة تركز على تنمية مهارات الملاحظة وطرح الأسئلة والتفسير العلمي، في استجابة للتوجه العالمي نحو تعزيز مهارات الاستقصاء العلمي في سن مبكرة ، ومع ذلك يواجه المعلمون تحديات عملية في تحقيق هذه الأهداف، ترتبط بأساليب التقويم التقليدية وضعف التدريب على استراتيجيات تعليم التفكير، وبشكل عام يمكن القول إن التطور التاريخي لتعليم مهارات التفكير العلمي اتجه من التركيز على المحتوى العلمي فقط إلى الاهتمام بتنمية عمليات العقل العلمية لدى المتعلمين، مدعومًا بنتائج البحوث التي تؤكد قابلية الأطفال لاكتساب مهارات كصياغة الفرضيات والتجريب المنهجي عند توفر بيئات تعلم مناسبة (Zimmerman, 2000; Kuhn, 2010).

ويُعدّ تعليم العلوم في المرحلة الابتدائية أحد أهم الأسس التي يقوم عليها بناء الشخصية العلمية للتلميذ؛ إذ تسهم المفاهيم والاستراتيجيات التي يتلقاها في هذه

الترقيم الدولي الموحد للطباعة ISSN: 3009-612X
الترقيم الدولي الموحد الالكتروني E. ISSN: 3009-6146
المجلة معرفة على قاعدة المجلات العلمية فى بنك المعرفة المصرى
وقاعدة بيانات Edu Search دار المنظومة

المرحلة في صقل مهاراته العقلية والمعرفية، كما تضعه على مسار فَهْم أعمق لما يجري في البيئة المحيطة به (Bransford, Brown, & Cocking, 2000). وفي إطار السعي نحو الارتقاء بالعملية التعليمية، شهدت المناهج الدراسية في العلوم تطورات مهمة تتضمن التركيز على الجانب العملي والاستقصائي، وتشجيع التلاميذ على طرح الأسئلة والتفكير الناقد وتجريب الفرضيات (NGSS, 2013) وعلى الرغم من هذه التطورات، لا تزال العديد من الطرق التقويمية التقليدية عاجزة عن توفير تغذية راجعة دقيقة وفورية حول مستوى فهم التلاميذ الحقيقي للمفاهيم العلمية، ولا سيما تلك المهارات التي تتطلّب تحليلاً أعمق وارتباطًا وثيقًا بالتجربة والملاحظة(Slavin, 2020) .

وُيقدم التقييم بالاختبارات الورقية أو بالأساليب التقليدية التي تعتمد على الدرجات النهائية صورة متجزأة عن أداء التلميذ (Hwang & Chang, 2011) فمثل هذه الأساليب لا تأخذ بالحسبان الفروق الفردية في التفكير العلمي، ولا تُلمّ بالعناصر الخفية في عملية التعلّم، مثل الوقت المستغرق لاستيعاب الفكرة، أو طبيعة الأخطاء المتكررة، أو توظيف التلميذ للمهارات العملية أثناء تنفيذ التجارب ، كما أنّ القدرة على جمع هذه البيانات وتحليلها يدوياً تشكّل عبئًا على المعلم، خاصة في الفصول كثيرة العدد؛ الأمر الذي يُضعف إمكانية التدخل المبكّر لمعالجة جوانب القصور العلمي لدى التلاميذ (Zawacki-Richter, Marín, Bond, & Gouverneur, 2019).

ومن هنا برز توجّه عالميّ نحو توظيف تقنيات الذكاء الاصطناعي وتعلُّم الآلة في التعليم، وبشكل خاص في جانب التقويم التربوي، وتندرج هذه الجهود تحت مجال "التنقيب في البيانات التعليمية (Educational Data Mining) وتحليلات التعلّم(Learning Analytics) ؛ حيث تهدف إلى استثمار وفرة البيانات المتراكمة عن أداء المتعلمين وخصائصهم، من أجل تطوير حلول ذكية توفّر صورة أكثر شمولًا وعمقًا حول مستويات الفهم والتحصيل الدراسي (Baker & Yacef, 2009; Romero & Ventura, 2020).

ويمكّن استخدام خوارزميات تعلم الآلة من التعرف على أنماط الأداء المعقدة، وتوقّع الصعوبات المستقبلية التي قد يواجهها بعض التلاميذ في مادة العلوم، مثل التعثر في فهم العلاقات السببية أو في تطبيق المفاهيم الفيزيائية البسيطة في تجارب المختبر.(Liu, Zhu, & Qin, 2021)

وبفضل التقدم الحاصل في تقنيات الحوسبة السحابية ومعالجة البيانات الضخمة، أصبحت عملية جمع البيانات التعليمية أسهل من ذي قبل، حيث يمكن استخدام التطبيقات التفاعلية وأجهزة الاستشعار وأدوات المحاكاة الافتراضية في بيئات التعلم؛ ما يساعد في توثيق كل خطوة يقوم بها التلميذ أثناء تنفيذ التجارب العلمية أو التفاعل مع الأنشطة الرقمية (Freeman et al., 2014). وبناءً على هذه البيانات الضخمة(Big Data) ، تتيح لنا خوارزميات التصنيف والتنبؤ واكتشاف الأنماط (Pattern Recognition) تحديد خصائص معيّنة للأداء

الترقيم الدولي الموحد للطباعة ISSN: 3009-612X
الترقيم الدولي الموحد الالكتروني E. ISSN: 3009-6146
المجلة معرفة على قاعدة المجلات العلمية فى بنك المعرفة المصرى
وقاعدة بيانات Edu Search دار المنظومة

العلمي، وتصنيف التلاميذ ضمن مستويات أو فئات معرفية مختلفة، مثل "مستوى متقدّم في التفكير الاستقصائي" أو "ضعيف في تطبيق مفهوم التغير الكيميائي"، ثم تقديم توصيات محددة تهدف إلى رفع مستوى الأداء لدى الفئة الضعيفة وتعزيز مهارات الفئة المتميزة.(Nguyen & Walker, 2016)

وفي مجال تصميم الأنشطة التعليمية الموجّهة بالبيانات (Data-Driven Education)، يمكننا الاستفادة من نتائج التحليل الآلي لتوجيه المناهج نحو جوانب الإثراء المناسبة ، فعندما يكشف النموذج التقويمي الذكي عن وجود ضعفٍ شائع بين مجموعة من التلاميذ في مهارة "توظيف الملاحظة التجريبية"، مثلاً، يمكن للمعلم تصميم أنشطة تفاعلية تركّز على تعزيز هذه المهارة، كإجراء تجارب علمية مبسّطة تتطلّب من التلاميذ توثيق ملاحظاتهم بصورة منهجية ومقارنة النتائج ، وبالتالي يصبح التقويم ليس مجرد رصد للدرجات وإصدار أحكام فقط ، بل حلقة متكاملة تغذّي العملية التعليمية وتطوّرها.(Clark & Feldon, 2014)

ويسهم استخدام الخوارزميات الذكية في توفير الوقت والجهد على المعلمين، إذ يمكن إعداد تقارير تحليلية متكاملة حول كل تلميذ –أو حول الفصل بأكمله– تشير إلى أنماط محددة تستدعي الانتباه، ويمكن لمثل هذه التقارير أن تتضمن رسومات بيانية وإحصاءات تُظهر تطوُّر فهم المفاهيم العلمية مع مرور الزمن، أو مقارنة تحصيل التلميذ بمجموعة أقرانه، أو حتى اقتراح قائمة بمهام تدريبية إضافية لكل تلميذ حسب نقاط ضعفه الفردية (Romero & Ventura, 2020).ولا يقتصر هذا الاهتمام على الجانب المعرفي وحسب، بل يشمل كذلك المهارات العلمية العملية؛ إذ يمكن أن ترصد النظم الذكية أداء التلميذ أثناء تنفيذ التجارب المخبرية، فتسجل جوانب مثل الالتزام بخطوات السلامة، واتباع تسلسل التجربة، والقدرة على تفسير النتائج بصورة صحيحة.(Hattie, 2009)

وأشار (Zawacki-Richter et al., 2019) إلى ضرورة دراسة التحديات المرتبطة بتطبيق هذه التقنيات في مدارس المرحلة الابتدائية؛ مثل الحاجة إلى توفير بنية تحتية تقنية (أجهزة لوحية، حواسيب، اتصال بالإنترنت) إلى جانب تدريب المعلمين على التعامل مع التقنيات الحديثة والتحليلات الرقمية، كما تطرح قضايا الخصوصية والأمان تساؤلات أخلاقية مهمة تتعلّق بمدى مأمونية البيانات التي يجمعها النظام الذكي، وضرورة حمايتها وفقًا لمعايير تشريعية تحدّ من إساءة الاستخدام وتكفل خصوصية التلاميذ .(Liu et al., 2021) وفي هذا الإطار، تؤكد منظمة الأمم المتحدة للتربية والعلم والثقافة (اليونسكو) على أهمية التوازن بين الاستفادة من تقنيات الذكاء الاصطناعي والالتزام بالمبادئ الإنسانية والأخلاقية في العملية التعليمية.(UNESCO, 2019)

وبناءً على ما سبق، هدف هذا البحث إلى تصميم نموذج تقويمي ذكي قائم على توظيف خوارزميات تعلم الآلة لتحليل البيانات التفصيلية حول أداء تلاميذ المرحلة الابتدائية في مادة العلوم، وتشخيص المهارات والمعارف التي قد تحتاج إلى تحسين أو تطوير. ويُطمح من خلاله إلى إبراز الكيفية التي يمكن أن يتحوّل

الترقيم الدولي الموحد للطباعة ISSN: 3009-612X
الترقيم الدولي الموحد الالكتروني E. ISSN: 3009-6146
المجلة معرفة على قاعدة المجلات العلمية فى بنك المعرفة المصرى
وقاعدة بيانات Edu Search دار المنظومة

فيها التقويم التربوي من كونه مجرد وسيلة لقياس التحصيل إلى أداةٍ تفاعليةٍ وديناميكيةٍ تعزّز من قدرات التلاميذ وتدعم نموهم العلمي بصورة مستدامة (Slavin, 2020). ومن المرتقب أن يُسهم هذا النظام في تمكين المعلمين من إعداد خطط علاجية وتطويرية أكثر فعالية، مستندة إلى أدلّة رقمية وتحليلات دقيقة، كما يتيح لصنّاع القرار التربوي والمتخصصين في تصميم المناهج فرصة اتخاذ قرارات قائمة على بيانات موثوقة، تكفل التحديث المستمر لطرق التدريس والتقويم في مجال العلوم.(Baker & Yacef, 2009)

وعليه، فإن الاستفادة من تكنولوجيا تعلُّم الآلة في تقويم العلوم تفتح آفاقًا جديدة نحو تحسين جودة التعليم في المراحل المبكرة، وتتيح إمكانية بناء أجيال تمتلك الأسس العلمية والمهارات التحليلية اللازمة لمواكبة التحولات المستقبلية في مختلف التخصصات العلمية والبحثية. كما تمثّل هذه المبادرة خطوة مهمة في مسار التعلم المخصص(Personalized Learning) ، حيث تُؤخذ بعين الاعتبار قدرات التلاميذ واهتماماتهم المختلفة، ويُصمّم المحتوى التدريبي والتقويمي بما يلائم احتياجاتهم الفردية. ومع التطوّر المستمر للخوارزميات والأدوات التقنية، تتعاظم فرص الربط بين الأنشطة العملية في الصفوف والمختبرات من جهة، والأدوات التحليلية الذكية من جهة أخرى، ليغدو التعلم عمليةً شاملة ومتكاملة، تستفيد من إنجازات التكنولوجيا في تمكين الطالب والمعلم على حدّ سواء (Freeman et al., 2014; Slavin, 2020).

مشكلة البحث:

شهد تعليم العلوم في المرحلة الابتدائية تحوّلاً من التركيز على تحصيل الحقائق إلى تنمية الكفايات العلمية القائمة على الاستقصاء وتحليل الأدلة ، ومع ذلك، ما تزال أدوات التقويم التقليدية—الاختبارات الورقية التحصيلية المعيارية والتصحيح اليدوي—تعتمد على منتَج التعلّم (الإجابة النهائية)، وتهمل مسار الأداء الذي يكشف خطوات التفكير العلمي(Zimmerman, 2007) وتشير تحليلات ما بعد الاختبار إلى أنّ هذا الأسلوب يُفضي إلى تشخيصٍ متأخِّر للفجوات المفاهيمية، ما يقلِّص فعالية التدخّلات العلاجية المبكرة (Slavin, 2020)

فعلى الرغم من الاهتمام المتزايد بتطوير مناهج العلوم في المراحل الدراسية المبكرة واعتماد استراتيجيات تدريس قائمة على البحث والاستقصاء والتجريب (NGSS, 2013)، لا تزال الأساليب التقويمية التقليدية تعاني من جوانب قصور عديدة، أبرزها التركيز على الاختبارات التحريرية والدرجات النهائية بوصفها المعيار الرئيس لقياس تحصيل التلميذ العلمي ، إن هذا التركيز الأحادي لا يعكس بالضرورة مستوى الفهم الحقيقي للمفاهيم أو مدى القدرة على تطبيق المهارات العلمية في مواقف حياتية أو تجريبية (Bransford, Brown, & Cocking, 2000)، فقد يتمكّن بعض التلاميذ من حفظ المعلومات نظريًا واجتياز الاختبار بنجاح، بينما يظلون غير قادرين على توظيف هذه المفاهيم

الترقيم الدولي الموحد للطباعة ISSN: 3009-612X
الترقيم الدولي الموحد الالكتروني E. ISSN: 3009-6146
المجلة معرفة على قاعدة المجلات العلمية فى بنك المعرفة المصرى
وقاعدة بيانات Edu Search دار المنظومة

بصورة عملية أو تفسير الظواهر الطبيعية التي تقابلهم في الحياة اليومية (Hwang & Chang, 2011).

ويتطلّب التدريس الفعال في العلوم التعرّف المبكّر على الصعوبات التي يواجهها التلميذ في استيعاب المفاهيم العلمية، مثل فهم العلاقات السببية والتفكير الاستنتاجي أو تحديد الفرضيات وتصميم التجارب البسيطة. إلا أنّ الأساليب التقليدية قد لا تسمح بتوفير هذه القراءة التفصيلية لأداء التلميذ العلمي؛ إذ عادةً ما يتم الاكتفاء بدرجة كلية تُعطى في نهاية الاختبار أو الفصل الدراسي (Romero & Ventura, 2020). وبناءً عليه، قد يفوت على المعلمين فرصة التدخل العلاجي المبكّر وتصحيح مسار التعلّم (Baker & Yacef, 2009). علاوةً على ذلك، تزداد الأمور تعقيدًا في الفصول ذات الأعداد الكبيرة من التلاميذ، حيث يترتب على المعلم أعباء عديدة في تحليل أداء كل تلميذ وتحليل أخطائه النوعية، الأمر الذي قد لا يتسنّى بصفة دقيقة ومتكررة (Zawacki-Richter, Marín, Bond, & Gouverneur, 2019).

وأشارت الدراسات الحديثة إلى أنّ حجم البيانات المتولدة عن عملية التعلّم –حتى في المرحلة الابتدائية– بات متنوعًا وكبيرًا، بفضل التطبيقات الرقمية والأنشطة التفاعلية وأدوات المختبر الافتراضي (Nguyen & Walker, 2016). ومع غياب نظامٍ يقدّم آليات تحليل شاملة لهذه البيانات، تُهدر فرصة توظيفها في الحصول على مؤشرات تنبؤية (Predictive Indicators) حول مستويات التعلّم ونوعية المهارات المكتسبة، وبخاصة مهارات التفكير العلمي العملي (Liu, Zhu, & Qin, 2021). وقد يترتب على ذلك فوات فرص حقيقية لتحسين العملية التعليمية وزيادة كفاءتها، حيث يظل المعلم يعتمد على خبرته الشخصية وتقديره الذاتي لاتخاذ قرارات تربوية مهمة مثل توجيه المحتوى الإضافي أو تعديل الأنشطة وفقًا لاحتياجات التلاميذ.

إلى جانب المشكلات التقنية والتربوية، تبرز أيضًا تحديات تتعلق بالفعالية الزمنية في إجراء التقويم وتقديم التغذية الراجعة؛ ففي كثير من الأحيان لا يتوفر للمعلم وقت كافٍ لإعداد اختبارات متكررة أو تحليل نتائجها بدقة، الأمر الذي ينعكس سلبًا على إمكانية إعطاء التلميذ توصيات فورية تضبط مسار تعلّمه وتدعمه (Hwang & Chang, 2011; Slavin, 2020)

ومما سبق يغدو التكامل بين التقويم التكويني المستمر وتقنيات الذكاء الاصطناعي خطوة مهمة نحو تعزيز جودة التعليم في مادة العلوم، إذ يستطيع نظامٌ ذكي يُوظّف خوارزميات تعلم الآلة جمع البيانات عن أداء التلميذ أثناء الأنشطة الصفية والتجارب العلمية، ثم تحليلها بصورة آنية لاستخلاص نقاط القوة والضعف لدى كل تلميذ على حدة (Romero & Ventura, 2020) وفضلًا عن ذلك، يمكن للنظام التنبؤ بمسار الأداء المستقبلي واقتراح تدخلات تربوية مخصصة من شأنها إيقاف تدهور أداء التلميذ قبل تفاقمه.(Baker & Yacef, 2009)

الترقيم الدولي الموحد للطباعة ISSN: 3009-612X
الترقيم الدولي الموحد الالكتروني E. ISSN: 3009-6146
المجلة معرفة على قاعدة المجلات العلمية فى بنك المعرفة المصرى
وقاعدة بيانات Edu Search دار المنظومة

لكن الواقع يشير أيضًا إلى عقبات ينبغي أخذها في الحسبان، كضرورة توفير بيئة تقنية ملائمة في المدارس، وتدريب المعلمين على تفعيل هذه الأدوات في الفصول الدراسية، والتعامل مع القضايا الأخلاقية والخصوصية(Zawacki-Richter et al., 2019). وعلى الرغم من صعوبة هذه التحديات، فإن الإمكانات الكامنة في حلول التعليم المبني على البيانات (Data-Driven Education) تبشّر بإحداث نقلات نوعية في مجال التقويم التربوي، إذا ما اقترنت بخطط واعية للتنفيذ والتدريب والمتابعة المستمرة (Hattie, 2009; Freeman et al., 2014)

بناءً على ما سبق، تتبلور مشكلة الدراسة في الحاجة الماسّة إلى نموذج تقويمي ذكي يستفيد من خوارزميات تعلم الآلة لتحليل البيانات التعليمية المتولدة عن التفاعلات الصفية والأنشطة العلمية لدى تلاميذ المرحلة الابتدائية، بما يكفل رصدًا مبكّرًا لنقاط الضعف في المفاهيم والمهارات العلمية، وتقديم توصيات تعليمية مخصصة لكل تلميذ، فضلًا عن إتاحة تقارير تحليلية للمعلم تسهّل مهامه اليومية في التقويم والمتابعة ، حيث تبرز هذه الحاجة كحل استراتيجي يهدف إلى رفع مستوى التحصيل في العلوم، وتعزيز الفهم الحقيقي للمفاهيم، وتحويل العملية التقويمية إلى أداة تشخيصية فعّالة تمهد الطريق للتطوير المستمر في أساليب التدريس، وبناء ثقافة علمية راسخة لدى الأجيال الناشئة(Slavin, 2020; Liu et al., 2021) .

ويستهدف نموذج التقويم الذكي توظّف خوارزميّات تعلّم الآلة لتحليل بيانات التفاعل الصفّي والأنشطة العلمية لدى تلاميذ الصف السادس الابتدائي، بهدف التشخيص المبكر لنقاط القوة والضعف في مهارات التفكير العلمي، الأمر الذي سيترتب عليه القدرة على توليد توصيات تعليمية مخصَّصة لكل تلميذ تستند إلى أداء لحظي موثوق، بدلًا من الدرجات التجميعية المتأخّرة، وإتاحة لوحات تقارير فورية للمعلّم تيسّر اتخاذ قرارات تدريس تفريدي وتخطيط أنشطة علاجية دقيقة.

ومجال تركيز الدراسة الحالية هو تصميم الخوارزمية، وتدريب النموذج، والتحقّق التجريبي من كفاءته (الدقّة، الاستدعاء، F-measure) باستخدام عيّنة بيانات واقعية من تفاعلات صفية في مدارس ابتدائية مصرية، أمّا بناء الواجهة الرسومية المتكاملة، وربط النظام ببيئة الصفّ الفعلية، ومعالجة قضايا النشر على نطاق واسع؛ فسوف يُتناول في أبحاث لاحقة بعد التحقق من صلاحية النموذج الأساسي.

وبهذا المسار المرحلي، تهدف الدراسة إلى سدّ فجوةٍ حاسمة بين إمكانات تقنيات التعلّم الآلي ومتطلبات التقويم التكويني في الممارسات الصفية، والمساهمة في تأسيس بنية تحتية بحثية وتعليمية تدعم الانتقال من التقويم الورقي المتأخّر إلى تقويم تشخيصي ذكي لحظي يرفع جودة تدريس العلوم ويعزّز الثقافة العلمية لدى الأجيال الناشئة.

الترقيم الدولي الموحد للطباعة ISSN: 3009-612X
الترقيم الدولي الموحد الالكتروني E. ISSN: 3009-6146
المجلة معرفة على قاعدة المجلات العلمية فى بنك المعرفة المصرى
وقاعدة بيانات Edu Search دار المنظومة

ومما سبق يمكن صياغة مشكلة الدراسة الحالية في السؤال الرئيس التالي:

ما مدى فاعليّة نظام تقييمٍ ذكي قائم على خوارزميّات تعلّم الآلة في تشخيص مستويات مهارات التفكير العلمي لدى تلاميذ الصف السادس الابتدائي، بالاعتماد على بيانات حقيقية مُجمَّعة من اختبار تحصيلي مقنَّن، ومهام أدائية، وقوائم ملاحظة للأداء العملي؟

أهداف البحث:

يهدف البحث الحالي إلى:

1. تصميم وتطوير نموذج تقويم ذكي يوظّف خوارزميّات تعلّم الآلة لتحويل مهارات للتفكير العلمي إلى تصنيف دقيق لمستوى التلميذ (متفوّق – متوسّط – ضعيف).
2. تدريب النموذج على بيانات حقيقيّة متعدّدة المصادر جُمعت بواسطة بطارية أدوات القياس التي تم تصميمها في الدراسة الحالية والتي اشتملت على (اختبار تحصيلي مقنَّن، مهام أدائية عمليّة، قوائم ملاحظة للأداء الصفّي)، وذلك لضمان شمول المقاييس وصدق التمثيل السلوكي للمهارات العلمية.
3. تقويم كفاءة النموذج إحصائيًا عبر مؤشرات الدقّة (Accuracy)، الاستدعاء (Recall)، الدقّة الإيجابية(Precision) ، ومعامل F1 ، ومقارنتها بطرائق التقويم التقليدية.

أهمية البحث:

أ- الأهمّيّة النظريّة:

- سدّ فجوة معرفيّة في مجال القياس التربوي العربي من خلال إرساء نموذج خوارزمي يربط بين ثماني مهارات أساسية للتفكير العلمي وبين تصنيف مستويات الأداء، وهو ما لم تُعالجه الأدبيّات المحليّة بصورة تجريبية منظَّمة حتى الآن.
- دمج مصادر بيانات متنوّعة (اختبار تحصيلي مقنَّن، مهام أدائية، قوائم ملاحظة) في إطار واحد للتعلّم الآلي؛ الأمر الذي يثري نماذج القياس النفسي بإدخال تمثيل سلوكي مباشر Not just outcome—scores—للمهارات العلمية.
- تقديم منهجية معيارية لتحويل درجات أدوات ورقية وأدائية إلى مدخلات رقمية قابلة للمعالجة الخوارزمية، بما يفتح الباب أمام باحثين آخرين لتطبيق المقاربة على مهارات أو مجالات معرفية مختلفة.

الترقيم الدولي الموحد للطباعة ISSN: 3009-612X
الترقيم الدولي الموحد الالكتروني E. ISSN: 3009-6146
المجلة معرفة على قاعدة المجلات العلمية فى بنك المعرفة المصرى
وقاعدة بيانات Edu Search دار المنظومة

ب- الأهمّيّة التطبيقيّة:

- تشخيص لحظي ودقيق لمستويات التفكير العلمي لدى تلاميذ الصف السادس، يسمح للمعلّم بالتدخّل المبكر ومعالجة القصور قبل أن تتحوّل الفجوات إلى صعوبات راسخة.
- تخصيص التعلم من خلال توليد تقارير تلقائية تُبرز نقاط القوة والضعف لكل متعلّم وتقترح مسارات علاجية أو إثرائية؛ ما يعزّز الدافعيّة ويرفع التحصيل العلمي بصورة مستدامة.
- ترشيد وقت وجهد المعلّم عبر الاعتماد على مؤشرات الأداء الآنيّة بدلاً من التصحيح اليدوي المطوَّل للمهام الأدائية وقوائم الملاحظة، فضلاً عن الحصول على تحليل إحصائي لا يتطلّب خبرة تقنية متقدمة.
- مقارنة موضوعيّة مع التقويم التقليدي تتيح لصنّاع القرار التربوي تقييم جدوى الاستثمار في أنظمة التقويم الذكية مقابل الأساليب الورقية، استناداً إلى مؤشرات دقيقة مثل الدقّة ومعامل F1.
- إسناد التحوّل الرقمي المدرسي برفع جاهزية المدارس الابتدائية لاستخدام تقنيات الذكاء الاصطناعي في التقويم، اتساقاً مع توجهات وزارة التربية والتعليم وخطط الدولة المصرية للتحوّل الرقمي.

الإطار النظري والدراسات السابقة:

ظهر مصطلح التفكير العلمي (Scientific Thinking)في أوائل القرن العشرين مع انتقال علم النفس من دراسة «الاستبصار» إلى دراسة عمليات الاستدلال على المشكلات الواقعية. وفي ثمانينيات القرن الماضي تبلور المفهوم داخل أبحاث تحليل حلّ المشكلات، ثم تطوّر في التسعينيات ليشمل البُعد الاجتماعي-الثقافي للعِلم (Koslowski, 1996) وقد اتفقت المراجعات الحديثة على أنّ التفكير العلمي ليس محتوى معرفيًا بقدر ما هو منظومة من العمليات المعرفية وما وراء المعرفية تمكّن الفرد من توليد معرفة قابلة للتفنيد.

ويُستخدم مصطلح التفكير العلمي على نطاق واسع في علم النفس المعرفي وعلوم التربية وفلسفة العلم، لكنّه يفتقر أحيانًا إلى تحديدٍ إجرائي متّفق عليه، تُشير (Koslowski, 1996) إلى أنّ معظم الالتباس يعود إلى الخلط بين "معرفة نتائج العلم" و"العمليات الذهنية التي تُنتِج تلك النتائج، وتُركّز معظم التعريفات على الوظائف العقليّة التي تدعم توليد الفرضيات، وفحص الأدلّة، وتعديل القناعات في ضوء البيّانات، وفيما يلي عرضاً لبعض تعريفات التفكير العلمي

وتُشير مهارات التفكير العلمي إلى مجموعة من المهارات العقلية والمنهجية التي يستخدمها الأطفال في هذه الفئة العمرية لفهم العالم من حولهم وحل المشكلات بطريقة منظمة تشبه أسلوب العلماء ، وهي شكل من السعي نحو المعرفة

الترقيم الدولي الموحد للطباعة ISSN: 3009-612X
الترقيم الدولي الموحد الالكتروني E. ISSN: 3009-6146
المجلة معرفة على قاعدة المجلات العلمية فى بنك المعرفة المصرى
وقاعدة بيانات Edu Search دار المنظومة

يمارسه معظم الناس بشكل هادف بهدف تحسين فهمهم وزيادة معرفتهم (Kuhn, 2002) ، كما تؤكّد تقارير التنمية المستدامة أنّ دولاً ذات نُظُم تعليمية متقدمة تُدرج التفكير العلمي ضمن مؤشرات الأداء القومي لجودة التعليم الأساسي ، ترى Kuhn (2002)أنّ التفكير العلمي هو "نشاط معرفي يهدف إلى بناء تفسيرات سببيّة عبر تنسيق الادّعاء مع الدليل، ومراقبة الفرد لعمليات تفكيره ذاتيًّا لضبط التحيّز واليقين" (Kuhn,2002) .

وتشمل هذه المهارات قدرات مثل: الاستقصاء وطرح الأسئلة والتفكير المنطقي المبني على الأدلة. عرف بعض الباحثين التفكير العلمي لدى الأطفال بأنه عملية منظمة لاكتساب المعرفة، تتضمن صياغة فرضيات واختبارها وجمع البيانات وتحليلها للوصول إلى استنتاجات مدعومة بالبراهين والأدلة التجريبية (Tobin & Capie, 1982).

وتصف Zimmerman (2007) مهارات التفكير العلمي بأنها المهارات المتضمنة في عمليات الاستقصاء والتجريب وتقييم الأدلة والاستدلال بهدف تطوير فهم الطفل للمفاهيم العلمية المحيطة به. أي أن الطفل الذي يمتلك هذه المهارات يكون قادرًا على اتباع منهجية علمية مبسطة في تفكيره اليومي، بحيث يلاحظ الظواهر، ويطرح الأسئلة حولها، ثم يجمع المعلومات والأدلة ليبني تفسيرًا منطقيًا لها. وتعد مهارات التفكير العلمي جزءًا أساسيًا من التفكير الناقد وحل المشكلات، إذ تساعد الأطفال ليس فقط في تعلم العلوم داخل المدرسة بل وفي التعامل مع مواقف الحياة اليومية بطريقة عقلانية قائمة على الأدلة بدلًا من التخمين العشوائي.

ويعرّفه Lawson (2000) على أنّه "مجموعة مهارات الاستدلال التي تشمل الاستقصاء، والتجريب، وتقييم الدليل، والاستنباط المنطقي؛ وهي مهارات تدعم تكوين المفاهيم العلمية وتطوير النظريات" (Lawson, 2000)

بحسب Lehrer & Schauble (2000) يُعَدّ التفكير العلمي "بناء نماذج تمثّل الظواهر الطبيعية وإعادة مواءمتها بصورة تكرارية لتحسين التنبّؤ ودقّة التفسير" (Lehrer & Schauble, 2000)

يعرّف تقرير PISA الصادر عن OECD (2019) Pisa التفكير العلمي بأنّه "القدرة على الانخراط مع قضايا ومسائل علمية بصفة مواطن، باستخدام الأدلة وتفسيرها نقديًّا لشرح الظواهر واتخاذ قرارات مستنيرة"

وقد عرّفت نماذج النموّ المعرفي التفكير العلمي بأنّه القدرة على (توليد تفسيرات سببيّة -تصميم تجارب مضبوطة -تفسير بيانات متغيّرة-**إصدار** أحكام مبنية على الأدلّة)، وتُشير مراجعة Kuhn (2011) إلى أنّ هذه المهارات تتطور على مسار متدرّج يبدأ بالملاحظة المجرّدة، وينتهي بالقدرة على تنسيق علاقات متعدّدة المتغيّرات، وفي السياق العربي، شدّد زيتون (2008) على أنّ كتب العلوم الحديثة يجب أن تُهيّئ بيئات صفّية تُمارَس فيها هذه العمليات بصورة يوميّة.

وتبيّن مجموعة من الدراسات الحديثة أن أطفال الصفوف من الرابع إلى السادس يُظهرون بالفعل مستويات أوّلية من التفكير العلمي إذا أتيحت لهم مهام

الترقيم الدولي الموحد للطباعة ISSN: 3009-612X
الترقيم الدولي الموحد الالكتروني E. ISSN: 3009-6146
المجلة معرفة على قاعدة المجلات العلمية فى بنك المعرفة المصرى
وقاعدة بيانات Edu Search دار المنظومة

استقصائية مناسبة ، فقد طوّر كل من Osterhaus, Koerber, Sodian, (2020) أداة **SPR-I** لقياس هذه المهارات، وأظهروا أنّ متوسط تلاميذ الصف الخامس يؤدّون أداءً يفوق التوقّعات التقليديّة في ثلاثة أبعاد: التجريب، وتفسير البيانات، وفهم طبيعة العلم (Osterhaus et al ., 2020) يدعم ذلك توصيات Zimmerman (2007) بأن القدرات الاستقصائيّة يمكن رصدها مبكرًا عبر أدوات قياس مُحكّمة.

ويوضح جدول (1) ثلاثة تعريفات مختلفة للتفكير العلمي تبعاً لاختلاف المدرسة النظرية

جدول (1): المدارس النظرية المختلفة وتعريفها للتفكير العلمي

المدرسة النظرية	التعريف المركزي	التركيز الأساسي
المعرفية-التجريبية (Zimmerman)	تنسيق فرضيات، تجارب، أدلّة، استنتاجات في دورة استقصاء مغلقة (Zimmerman, 2007)	البنية الإجرائية
الأبستمولوجيا /المعرفية (Kuhn)	نشاط معرفي هدفه بناء تفسيرات سببية عبر مناقشة الأدلة ومراقبة التفكير ذاتيًا (Kuhn, 2002)	التقييم النقدي للأدلة
النمذجة (Lehrer & Schauble)	استخدام تمثيلات رمزية أو مادية لوصف الظواهر وتوليد تنبؤات قابلة للاختبار (Lehrer & Schauble, 2000)	بناء النماذج وتعديلها

ورغم الاختلاف بين هذه التعريفات إلا أنها تشترك في أربع سمات جوهرية :السببية، الاختبار التجريبي، استناد النتائج للأدلّة، ومراجعة القناعة.

ويسعى علماء النفس والتربية إلى فهم كيف تتطور مهارات التفكير العلمي لدى الأطفال، وتم اقتراح عدة نماذج نظرية تفسر بنية هذه المهارات ومسار نموها عبر مراحل الطفولة، من أبرز هذه النماذج:

وفقًا لبياجيه، يمر الأطفال في عمر 6–12 سنة بمرحلة العمليات العيانية (الملموسة) يليها بدايات مرحلة العمليات الصورية (التجريدية) قرب سن 11–12 عامًا(Inhelder & Piaget, 1958) ، في المرحلة الملموسة يكون تفكير الطفل مرتبطًا بالمحسوسات ويتمكن من إجراء عمليات عقلية أساسية على الأشياء المادية التي يراها أو يجربها مباشرة، أما في مرحلة العمليات الصورية، فيبدأ الطفل باكتساب القدرة على التفكير المجرد وافتراض فرضيات ذهنية دون الحاجة إلى معايشة فعلية لكل حالة.

هذا الانتقال مهم للتفكير العلمي، إذ أنه في عمر (11) تقريبًا يصبح العديد من الأطفال قادرين على التفكير الفرضي الاستنتاجي؛ أي تخيل فرضيات متعددة حول مشكلة ما والتفكير بمنطق "ماذا لو" واستخلاص نتائج منطقية منها حتى قبل إجراء التجربة فعليًا ، فعلى سبيل المثال، أشار بياجيه إلى تجربة البندول التي أوضحت كيف أن الأطفال الأصغر سنًا يغيرون عدة عوامل معًا دون منهجية، بينما يتمكن المراهقون الأكبر (حوالي 12 سنة فأكثر) من التحكم المنهجي بالمتغيرات (كطول الخيط وثقل الثقل) لمعرفة تأثير كل منها على حدة في تأرجح البندول (Inhelder & Piaget, 1958) .

الترقيم الدولي الموحد للطباعة ISSN: 3009-612X
الترقيم الدولي الموحد الالكتروني E. ISSN: 3009-6146
المجلة معرفة على قاعدة المجلات العلمية فى بنك المعرفة المصرى
وقاعدة بيانات Edu Search دار المنظومة

ويُشير "بياجيه" إلى أن نضج القدرات العقلية لدى الطفل (بحكم العمر والتفاعل مع البيئة) هو ما يسمح بالتطور التدريجي في مهارات التفكير العلمي، حيث يصل الطفل في النهاية إلى مستوى يستطيع معه التفكير كالراشدين ولكن بشكل مبسط.

وفي المقابل يؤكد النموذج الاجتماعي الثقافي لـ Vygotsky على دور العوامل الاجتماعية والتفاعل مع الكبار في تطور التفكير العلمي، حيث يُشير إلى أن تعلم الطفل يحدث ضمن منطقة النمو الأقرب لديه عبر التفاعل مع أشخاص أكثر خبرة (كالوالدين والمعلمين) واللغة التي يتعلم من خلالها المفاهيم والأفكار العلمية (Vygotsky, 1978)

ويُضيف أن مهارات التفكير العلمي لدى الطفل ليست نتاج النضج البيولوجي وحده، بل هي أيضًا نتاج التعليم الموجه والدعم الاجتماعي، فعلى سبيل المثال، قد لا يكتشف الطفل من تلقاء نفسه كيفية إجراء تجربة مضبوطة المتغيرات، لكن بإرشاد المعلم وتوجيه الأسئلة المناسبة يستطيع الطفل تعلم أسلوب التجريب المنهجي حتى قبل أن يصل إلى مرحلة العمليات الصورية من حيث التطور العقلي.

كما أكد على التمييز بين "المفاهيم العفوية "التي يكوّنها الطفل من خلال خبراته اليومية وبين "المفاهيم العلمية "التي يكتسبها من خلال التعليم المنظم في المدرسة ، اكتساب المفاهيم العلمية (مثل مفهوم القوة أو الكتلة أو منهجية التجريب العلمي) ، يُنمِّي لدى الطفل أساليب تفكير أكثر تجريدًا ونظامًا مقارنةً بمفاهيمه العفوية البسيطة ، وبذلك توفر البيئة التعليمية الداعمة فرصًا للطفل لتطوير مهارات التفكير العلمي من خلال النقاش والشرح وحل المشكلات بشكل تفاعلي ضمن مجتمع الصف – حيث يلعب المعلم دور الميسر الذي ينقل الطفل إلى مستويات عليا من التفكير العلمي تقع ضمن نطاق إمكانياته القريبة.(Vygotsky, 1978)

وقد قدمت نماذج معرفية معاصرة رؤى إضافية حول كيفية تطور مهارات التفكير العلمي ، أحد هذه التوجهات هو اعتبار الطفل "عالِمًا صغيرًا " بالمعنى الحرفي، حيث يقترح علماء مثل Gopnik (2012) أن الأطفال منذ سن مبكرة يقومون ببناء نظريات صغيرة عن العالم من حولهم، ثم يقومون باختبار هذه النظريات وتعديلها بناءً على ما يواجهون من أدلة، حيث يندمج الأطفال في عمليات مشابهة للاستقصاء العلمي أثناء لعبهم واستكشافهم العفوي ، على سبيل المثال، الطفل الذي يجرب طرقًا مختلفة لتركيب قطع البناء ليكتشف أيها أكثر ثباتًا يقوم فعليًا بصياغة فرضيات ضمنية واختبارها واحدًا تلو الآخر. يشير Gopnik (2012)إلى أن لدى الأطفال الصغار قدرات استدلال علمي أولية تسمح لهم باكتشاف علاقات سببية من خلال الملاحظة والتجربة في إطار اللعب الحر.

هذا النموذج المعروف أحيانًا باسم "نظرية العقل العلمي" لدى الأطفال يؤكد أن أسس التفكير العلمي موجودة فطريًا وتتطور مع التجربة، وليس الأمر أن الطفل يفتقر تمامًا لتلك القدرات قبل مرحلة عمرية معينة.

الترقيم الدولي الموحد للطباعة ISSN: 3009-612X
الترقيم الدولي الموحد الالكتروني E. ISSN: 3009-6146
المجلة معرفة على قاعدة المجلات العلمية فى بنك المعرفة المصرى
وقاعدة بيانات Edu Search دار المنظومة

من جهة أخرى، طوّر باحثون في علم النفس المعرفي نماذج تركز على استراتيجيات التفكير العلمي وكيفية تعلمها، فقد اقترح كل من Klahr & Dunbar (1988)نماذج لوصف عملية حل المشكلات العلمية مثل نموذج البحث الثنائي الذي يصور قيام الفرد بالبحث في فضاء الفرضيات وفضاء التجارب بشكل متوازٍ عند محاولة اكتشاف الحل .(Klahr & Dunbar, 1988) وضمن الدراسات التجريبية، تبين أنه يمكن تعليم الأطفال استراتيجيات محددة للتفكير العلمي في سن مبكرة نسبيًا.

فقد أثبتت دراسة Chen and Klahr (1999) أن أطفالًا في الصف الثاني والرابع الابتدائي (أعمار 7–10 سنوات) استطاعوا تعلم وإتقان استراتيجية التحكم في المتغيرات (أي تصميم تجارب يغيّرون فيها متغيرًا واحدًا في كل مرة) بعد تلقي تعليمات وتدريبات موجهة. تمكن هؤلاء الأطفال – بمساعدة منهجية تدريسية مناسبة – من تصميم تجارب مضبوطة واستخلاص استنتاجات صحيحة منها، وهي مهارة كان بياجيه يعتقد أنها لا تتبلور بشكل كامل قبل مرحلة المراهقة المتأخرة.

يتفق ذلك مع الاتجاه النظري الذي يؤكد على دور التدريب والتعلم الموجه في صقل مهارات التفكير العلمي، حيث يمكن تسريع تطور بعض المهارات أو تعزيزها عبر التدخل التعليمي حتى وإن لم يكن الطفل قد وصل بعد إلى مرحلة النمو المعرفي التي توقعها بياجيه لهذه المهارات ، كما أشار عدد من الباحثين المعاصرين إلى أهمية تنمية القدرات ما وراء المعرفية لدى الأطفال – مثل وعيهم بطرق تفكيرهم ومراقبتهم لصحة استدلالهم – كجزء من تطوير التفكير العلمي المنهجي .(Kuhn & Pearsall, 2000) ، فكلما تقدم الأطفال في العمر واكتسبوا خبرات أكثر، يصبحون أقدر على تقييم أفكارهم الخاصة ونقدها (مثلاً يدرك الطفل أن تجربته قد تحتوي على متغيرات مشوشة تؤثر على النتيجة) وهذا الوعي بالعملية العلمية يعزز نضج مهارات التفكير العلمي بشكل أكبر.

بشكل عام، تجمع النماذج النظرية المختلفة على أن تطور مهارات التفكير العلمي عملية معقدة تتأثر بكل من نمو الطفل العقلي الذاتي (النضج والتطور البيولوجي) وبالخبرات التعليمية والاجتماعية المحيطة به، فبينما يوفر النمو المعرفي الأساس للقدرات (مثل الذاكرة العاملة والمنطق)، تقوم التفاعلات التعليمية والاجتماعية بتوجيه هذه القدرات وصقلها نحو أساليب تفكير علمية محددة. وعليه، فإن الطفل في سن 6–12 سنة ينتقل تدريجيًا من التفكير الملموس نحو التفكير الأكثر تجريدًا، لكن مقدار ما يطوره من مهارات علمية في هذه الرحلة يعتمد كثيرًا على البيئة التعليمية والتدريبية التي يتعرض لها بالإضافة إلى تجاربه الذاتية في الاستكشاف والفضول العلمي الفطري.

الترقيم الدولي الموحد للطباعة ISSN: 3009-612X
الترقيم الدولي الموحد الالكتروني E. ISSN: 3009-6146
المجلة معرفة على قاعدة المجلات العلمية فى بنك المعرفة المصرى
وقاعدة بيانات Edu Search دار المنظومة

مراحل نمو وتطور التفكير العلمي عند الأطفال:

- **الانتقال من الحدس إلى السببية المنظمة:**

اعتمدت النظريات الكلاسيكية على تسلسل بياجيه، غير أنّ الأدلة الحديثة تُظهر أنّ الأطفال قادرون على أشكال مصغّرة من الاستدلال العلمي قبل المرحلة الصورية، فدراسة Gopnik (2012)تبرهن أنّ أطفال أربع سنوات يختبرون فرضيات مبنية على نوع من أنواع التعلّم الاحتمالي (Gopnik, 2012)، ويوضح جدول (2) مراحل تطور التفكير العلمي لدى الأطفال.

جدول (2) مراحل تطور التفكير العلمي لدى الأطفال

المرحلة العمرية بالسنوات	ملامح النمو الرئيسة
3–5	ملاحظة حسّية وتمييز سببي بدائي؛ استخدام تجارب «لو…فإن» في اللعب دراسات اللعب الافتراضي وموازين الضوء-الصوت (Gopnik, 2012)
6–8	ظهور CVS (السبب والنتيجة) لتفسير أنماط ثنائية، ويمكن ذلك عندما تُبسّط المهام مثال على العلاقة بين (رمل/منحدر)؛ (Chen & Klahr, 1999; Klahr & Nigam, 2004)
9–12	تكامل البيانات مع التوقع النظري؛ فهم عدم اليقين الإحصائي الأولي Koerber et al., 2015
>12	تفسير أحادي ومتعدد المتغيّرات، ووعي بطبيعة العلم (Zimmerman, 2007)

- **اكتساب القدرة على معرفة مفهوم السبب والنتيجة:**

يُظهر بحث Chen & Klahr (1999) أنّ تدريبًا صريحًا لثلاث جلسات كافٍ لرفع استخدام مفهوم السبب والنتيجة من 25 ٪ إلى 70 ٪ لدى تلاميذ 7-10 سنوات، وأن الأثر ينتقل إلى سياقات جديدة بعد أسبوعين (Chen & Klahr, 1999)والتفسيرات اللفظية للتجربة عزّزت الفهم المنطقي، لا الإجراء الميكانيكي فقط.

- **تطور تفسير البيانات:**

أظهر Masnick & Morris (2008) أنّ الأطفال في الصف الرابع يركزون على الاتجاه العام ويغفلون عن التشتت، بينما يحلول الصف السادس يبدؤون في دمج الحجم والعينات المتباينة في أحكامهم (Masnick & Morris, 2008).

الترقيم الدولي الموحد للطباعة ISSN: 3009-612X
الترقيم الدولي الموحد الالكتروني E. ISSN: 3009-6146
المجلة معرفة على قاعدة المجلات العلمية فى بنك المعرفة المصرى
وقاعدة بيانات Edu Search دار المنظومة

- **ظهور النمذجة السببية:**

بيّن Lehrer & Schauble أنّ تلاميذ الصف الأول قادرون على تمثيل ظاهرة واحدة في نماذج مادية، لكنهم لا ينتقلون إلى النمذجة الرياضية إلا بعد أن يفهموا فكرة "المتغيّر غير المرصود" في الصف الثالث تقريبًا (Lehrer & Schauble, 2000).

- **عوامل مؤثرة في مسار التطور:**

1. المعرفة المسبقة: الثراء المفاهيمي يُسهّل توليد فرضيات ذات جدوى (Koslowski, 1996).
2. ما وراء المعرفة: قدرة الطفل على مراقبة الخطأ والتعلّم منه ترتبط بنضج الفصّ الجبهي الأوسط، وتتحسن بتوجيهات التقييم الذاتي (Kuhn, 2002)
3. التعليم الصريح مقابل الاكتشاف الموجّه: التحليلات البعدية تُظهر أن أفضل النتائج تتحقق بمزيج من التوجيه والفرصة الحرة للتجريب (Klahr & Nigam, 2004)

وتُشير نتائج الدراسة التي قام بها كل من Achiaw & Owusu (2023) أنّ بيئات منخفضة الموارد تعيق تطور CVS بنسبة تصل إلى 30 ٪ مقارنة بالمدارس الحضرية، وأن توفير أدوات بسيطة (ماء، رمل، قياس شريط) يقلّص الفجوة إلى النصف خلال فصل دراسي(Achiaw & Owusu, 2023)

توصي دراسات تقدّم التعلم (Learning Progressions) بوصف مراحل متتابعة قابلة للقياس تبدأ من "التفسير اليومي" وتنتهي بـ "التفسير النظري" (Liu & Schwartz, 2016) ، وهذا يقتضي أدوات تشخيص تحاكي السياقات اليومية وتتدرّج في التعقيد دون الانتقال المفاجئ إلى صيغ رمزية مجردة.

فالتفكير العلمي يظهر مبكرًا كوميض حدسي، ثم يُصقل بتراكم الخبرات والتعليم المقصود. تكشف الأدلة أنّ:

- المعارف المسبقة والمهارات ما وراء المعرفية تعملان كعوامل تحفيز وتشريع لتطور مهارات التفكير العلمي لدى الأطفال.
- التدخلات التي تتبع منطق التدرّج الهرمي للمهارات تحقق نتائج أكبر.
- التكييف الثقافي للأدوات ضروري لضمان عدالة القياس.

هذه النتائج تشدّد على الحاجة إلى تصميم أدوات تشخيصية تراعي الخصائص النمائية لكل فئة عمرية، وتستخدم سياقات غنيّة ومحفّزة كي ترصد النمو الحقيقي لا الأداء الشكلي.

الترقيم الدولي الموحد للطباعة ISSN: 3009-612X
الترقيم الدولي الموحد الالكتروني E. ISSN: 3009-6146
المجلة معرفة على قاعدة المجلات العلمية فى بنك المعرفة المصرى
وقاعدة بيانات Edu Search دار المنظومة

الأبعاد الرئيسية لمهارات التفكير العلمي:

يمكن تحليل مهارات التفكير العلمي لدى الأطفال إلى عدد من الأبعاد أو المكونات الأساسية التي تميز الطريقة العلمية في التفكير، يقسم العديد من الباحثين هذه المهارات إلى فئات فرعية لتسهيل دراستها وتعليمها – أحيانًا تحت مسمى "مهارات العمليات العلمية" – حيث تُصنَّف إلى مهارات أساسية بسيطة وأخرى مركبة أكثر تقدمًا .(Padilla, 1990; Tobin & Capie, 1982) فيما يلي أبرز المهارات الفرعية التي تتضمنها عملية التفكير العلمي لدى الأطفال من عمر 6 إلى 12 سنة:

- الملاحظة :(Observation) قدرة الطفل على استخدام الحواس لجمع المعلومات عن الظواهر والأشياء بدقة وانتباه. تشمل الملاحظة تدوين السمات والخصائص وتحديد أوجه الشبه والاختلاف بين الأشياء. وتعد الملاحظة نقطة البداية لأي استقصاء علمي ناجح، فهي التي تزود الطفل بالمعطيات الأولية عن البيئة المحيطة به.
- التصنيف :(Classification) قدرة الطفل على تنظيم المعلومات من خلال تجميع الأشياء أو الكائنات ضمن فئات بناءً على خصائص مشتركة. يساعد التصنيف على تبسيط الكم الكبير من المعلومات بتقسيمها إلى مجموعات ذات معنى، مما يسهل على الطفل فهم العلاقات بين العناصر المختلفة. على سبيل المثال، قد يقوم الطفل بفرز المواد إلى فئتين (مواد تجذب المغناطيس وأخرى لا تجذب) بعد عدة ملاحظات.
- القياس :(Measurement) قدرة الطفل على استخدام أدوات ووحدات قياس مناسبة لتكميم الخصائص والمقادير (مثل الطول، الكتلة، الحجم، الزمن)، فالقياس الدقيق عنصر أساسي في التحقق العلمي لأنه يمكن الطفل من مقارنة النتائج بشكل موضوعي. يتعلم الأطفال في هذه المرحلة استخدام أدوات بسيطة كالمسطرة والميزان والساعة لإجراء قياسات تدعم ملاحظاتهم.
- الاستنتاج :(Inference) مهارة استخلاص النتائج أو التفسيرات من الملاحظات والبيانات المتوفرة. يعتمد الاستنتاج على التفكير المنطقي بحيث يربط الطفل بين ما تمّت ملاحظته وما يمكن أن يعنيه ذلك بالنسبة لسؤاله أو مشكلته، مثلًا، قد يستنتج الطفل أن النبات الذي أوراقه صفراء لم يحصل على ماء كافٍ بناءً على ملاحظاته السابقة عن حاجة النباتات للماء.
- التنبؤ :(Prediction) قدرة الطفل على توقع ما قد يحدث في المستقبل بناءً على أنماط أو اتجاهات ملحوظة في البيانات أو الخبرات السابقة. يستخدم التنبؤ المعرفة المكتسبة والاستدلال للاستشراف بما سيحدث إذا استمرت نفس الظروف أو تم تغيير متغير ما. على سبيل المثال، يمكن للطفل أن يتنبأ بنتيجة

الترقيم الدولي الموحد للطباعة ISSN: 3009-612X
الترقيم الدولي الموحد الالكتروني E. ISSN: 3009-6146
المجلة معرفة على قاعدة المجلات العلمية فى بنك المعرفة المصرى
وقاعدة بيانات Edu Search دار المنظومة

تجربة علمية (مثل توقع اتجاه نمو النبات إذا وُضع في مكان مظلم) بالاعتماد على معلوماته عن حاجات النبات للضوء.

- صياغة الفرضيات (Formulating Hypotheses): قدرة الطفل على اقتراح تفسيرات مؤقتة أو توقعات مبررة يمكن اختبارها. الفرضية عبارة عن تخمين مستنير حول سبب ظاهرة أو حل لمشكلة، يصيغه الطفل بناءً على معرفته وملاحظاته. تعكس هذه المهارة التفكير "ماذا لو"، حيث يتخيل الطفل تفسيرًا ممكنًا ثم يفكر في طريقة لفحص صحة هذا التفسير.
- التجريب (Experimentation): وتشمل التخطيط للبحث واختبار الفرضيات عمليًا .تتطلب هذه المهارة أن يصمم الطفل تجربة بسيطة للتحقق من فرضيته، وذلك عبر تحديد المتغيرات (مستقلة، تابعة، وضابطة) والتحكم فيها قدر الإمكان، وتنفيذ خطوات التجربة بطريقة منظمة وآمنة. خلال التجريب، يجمع الطفل البيانات والمشاهدات المرتبطة بسؤاله. وتمثل مهارة التجريب ذروة التفكير العلمي حيث ينتقل الطفل من مرحلة التفكير النظري إلى اختبار الأفكار عمليًا.
- تفسير البيانات ونتائج التجربة (Data Interpretation): قدرة الطفل على تحليل النتائج التي حصل عليها من الملاحظات أو التجارب، ومحاولة فهم ماذا تعني فيما يتعلق بالسؤال المطروح أو الفرضية المفترضة. يتضمن ذلك تنظيم البيانات (مثلاً عبر رسوم بيانية أو جداول مبسطة)، والبحث عن أنماط أو علاقات سببية، ثم استخلاص استنتاجات تدعمها هذه الأدلة. قد يكتشف الطفل من خلال تفسير البيانات أن فرضيته كانت صحيحة أو خاطئة، أو ربما بحاجة إلى تعديل.
- التواصل العلمي (Scientific Communication): وهي مهارة نقل المعلومات والأفكار العلمية للآخرين بوضوح ودقة. تشمل قدرة الطفل على وصف ما قام به من إجراءات وما توصل إليه من نتائج، سواء شفهيًا أو كتابيًا أو من خلال رسوم توضيحية. التواصل جزء مهم من العملية العلمية لأن مشاركة النتائج وتفسيرها تتيح للآخرين فهم عمل الطفل وإعادة تجربته أو البناء عليه. يتعلم الأطفال في سن المدرسة الأساسية التعبير عن أفكارهم العلمية بجمل واضحة وتقديم مبررات تدعمها الأدلة التي جمعوها.

تشكّل هذه الأبعاد مجتمعة بنية مهارات التفكير العلمي لدى الطفل؛ فهي مترابطة وتعزز بعضها بعضًا، على سبيل المثال، الملاحظة الدقيقة تقود إلى تصنيف أفضل للمعلومات، مما يساعد في صياغة فرضيات أكثر دقة، ويمكن اختبار هذه الفرضيات عبر التجريب وجمع البيانات ثم تفسيرها للوصول إلى استنتاجات واضحة يتم التواصل بشأنها. وقد أكدت الأدبيات التربوية والعلمية

الترقيم الدولي الموحد للطباعة ISSN: 3009-612X
الترقيم الدولي الموحد الالكتروني E. ISSN: 3009-6146
المجلة معرفة على قاعدة المجلات العلمية فى بنك المعرفة المصرى
وقاعدة بيانات Edu Search دار المنظومة

أهمية إتقان الأطفال لهذه المهارات، حيث تُعد حجر الأساس لفهمهم مناهج العلوم وتكوين عقلية علمية ناقدة منذ الصغر (Padilla, 1990; Zimmerman, 2007).الجدير بالذكر أن المهارات العلمية الأساسية (مثل الملاحظة والتصنيف) تبدأ في التشكل في سن مبكرة من المرحلة الابتدائية، في حين أن المهارات المركبة (مثل التحكم بالمتغيرات وصياغة الفرضيات) تصبح أكثر وضوحًا في نهاية المرحلة الابتدائية وبداية المتوسطة مع نمو قدرات الطفل العقلية (Tobin & Capie, 1982).وهذا التدرج يتوافق مع التطور المعرفي العام للطفل، حيث تبنى المهارات البسيطة الأساس للمهارات الأكثر تعقيدًا.

أدوات وطرق قياس مهارات التفكير العلمي عند الأطفال:

لقياس مدى اكتساب الأطفال لمهارات التفكير العلمي، ابتكر الباحثون والتربويون عدة أساليب وأدوات تقييم تجمع بين الاختبارات المقننة والملاحظات النوعية للأداء، يعد تقييم هذه المهارات أمرًا مهمًا وصعبًا في الوقت ذاته؛ فالمهم هو الحصول على صورة دقيقة لقدرات الطفل العلمية بعيدًا عن تأثير مهارات أخرى مثل القراءة أو الحفظ، فيما يلي أبرز الطرق المستخدمة لتقييم مهارات التفكير العلمي للأطفال في مرحلة 6–12 سنة:

1- الاختبارات القياسية (Standardized Tests): طُورت اختبارات ورقية-قلمية معيارية لقياس مهارات التفكير العلمي بشكل كمي موضوعي ، على سبيل المثال، قام Tobin & Capie (1982) بوضع اختبار للمهارات العلمية المتكاملة يقيس قدرة الطالب على تحديد المشكلات وفرض الفرضيات والتحكم في المتغيرات وتفسير البيانات من خلال أسئلة اختيار من متعدد وأخرى تتطلب ترتيب خطوات التجربة ، كما صمم Lawson (1978)اختبارًا للتفكير العلمي المنطقي (يعرف باختبار لوسون) يهدف إلى قياس مستوى تطور تفكير الطالب وفق مفاهيم بياجيه، بما في ذلك مهارات مثل استنباط العلاقات والتحليل المنطقي والتحكم بالمتغيرات.

تتميز هذه الاختبارات بأن لها معايير معيارية (Norms) يمكن من خلالها مقارنة أداء الطفل بأقرانه في نفس المرحلة العمرية ، ويتم حساب درجات تعبر عن مستوى إتقان الطفل لكل مهارة من المهارات العلمية ، على سبيل المثال، قد يكشف اختبار مقنن أن طالبًا معينًا يجد صعوبة في تحديد فرضية مناسبة منطقيا لمشكلة معينة، أو أنه لم يطور بعد فهمًا كافيًا لفكرة المتغيرات الضابطة في التجربة. مثل هذه المعلومات تفيد المعلمين والباحثين في تقييم جوانب القوة والضعف في تفكير الطفل العلمي. وتجدر الإشارة إلى أن الاختبارات القياسية توفر موضوعية وسهولة في التصحيح، لكنها قد لا تغطي سوى جانب من المهارات وقد تعجز عن قياس مهارات الأداء العملي الفعلي للطفل في سياق حقيقي.

2- الملاحظة العملية والأداء العملي (Performance Assessment): إلى جانب الاختبارات المكتوبة، تعتمد طرق تقييم أخرى على ملاحظة أداء

الترقيم الدولي الموحد للطباعة ISSN: 3009-612X
الترقيم الدولي الموحد الالكتروني E. ISSN: 3009-6146
المجلة معرفة على قاعدة المجلات العلمية فى بنك المعرفة المصرى
وقاعدة بيانات Edu Search دار المنظومة

الطفل أثناء قيامه بنشاط علمي فعلي .في هذا النوع من التقييم، يُطلب من الطفل إنجاز مهمة علمية (كتنفيذ تجربة بسيطة أو حل مشكلة عبر خطوات منهجية) ويتم رصد سلوكه وتفكيره خلال المهمة .قد يقوم المعلم أو الباحث بإعطاء الطفل أدوات وتجهيزات لإجراء تجربة معينة – مثل مجموعة من المواد لاكتشاف أيها يذوب في الماء – ثم يراقب كيفية تنظيم الطفل للتجربة: هل يضع خطة؟ هل يغير متغيرًا واحدًا في كل مرة أم عدة متغيرات معًا؟ هل يدون ملاحظاته؟ وكيف يستخلص النتائج؟ بعد ذلك يتم تقييم المهارات التي استخدمها الطفل وفق قائمة تحقق أو Rubricمحدد.

هذه الطريقة تسمح برؤية مباشرة لكيفية تطبيق الطفل للمهارات في موقف عملي حقيقي، وبالتالي توفر فهمًا أعمق لجودة تفكيره العلمي .وقد استخدمت العديد من الدراسات التعليمية هذا الأسلوب لتقييم اكتساب المهارات العلمية، حيث يُمنح الطالب مهمة مفتوحة النهايات نسبيًا ويُلاحظ أداؤه من حيث استخدام المهارات المختلفة .(Harlen, 1999) على سبيل المثال، أظهرت أساليب التقييم بالأداء أن بعض الأطفال قد يعرفون نظريًا خطوات الطريقة العلمية ولكن يواجهون صعوبة في تنفيذها عمليًا بشكل منظم عندما يُتركون دون إرشاد مباشر.

3- المقابلات الفردية والأسئلة المفتوحة: تعتبر المقابلة الإكلينيكية التي ابتكرها بياجيه من أقدم طرق تقييم تفكير الأطفال، ولا تزال تُستخدم بتعديلاتها الحديثة في مجال أبحاث التفكير العلمي. في هذه الطريقة، يجلس الباحث مع الطفل بشكل فردي ويطرح عليه أسئلة نوعية تستهدف الكشف عن فهمه لمفاهيم علمية معينة أو استدلاله في موقف افتراضي. قد يعرض الباحث على الطفل مشكلة افتراضية أو تجربة فكرية ثم يسأله عما يعتقد أنه سيحدث ولماذا .على سبيل المثال، يمكن أن يُعرض على الطفل رسم لتجربة فيها أكثر من متغير تم تغييره معًا، ويُسأل: "هل نستطيع من هذه التجربة أن نعرف أي عامل تسبب في النتيجة؟ ولماذا؟".

هذا النوع من الأسئلة يتطلب من الطفل أن يميز بين الفرضية والدليل وأن يفهم مفهوم التجربة المضبوطة، وقد استخدم Sodian et al. (1991) مثل هذه المقابلات لدراسة قدرة الأطفال (بعمر 7–10 سنوات) على التفريق بين الفرضيات والأدلة، حيث طُلب منهم الحكم على تجارب وهمية فيما إذا كانت نتائجها تدعم فرضية معينة أو لا.

أظهرت تلك الدراسة أن كثيرًا من الأطفال في نهاية الابتدائية يستطيعون إدراك أن التجربة غير المحكمة (التي تتغير فيها عدة عوامل معًا) لا تسمح باستنتاج سببي واضح، مما يدل على تطور فهمهم لمنهجية اختبار الفرضيات (Sodian, Zaitchik, & Carey, 1991)توفر المقابلات والمناقشات المفتوحة رؤية تفصيلية لطريقة تفكير الطفل ومنطقه، إذ يمكن للباحث تتبع خطوات تفكير

الترقيم الدولي الموحد للطباعة ISSN: 3009-612X
الترقيم الدولي الموحد الالكتروني E. ISSN: 3009-6146
المجلة معرفة على قاعدة المجلات العلمية فى بنك المعرفة المصرى
وقاعدة بيانات Edu Search دار المنظومة

الطفل وسؤاله عن تبريراته عند كل خطوة. ورغم أن هذه الطريقة نوعية وتستهلك وقتًا وجهدًا كبيرًا، لكنها غنية بالمعلومات وتساعد في الكشف عن سوء الفهم أو التصورات الخاطئة التي قد لا تظهر في الاختبارات التقليدية.

4- المهمات والمشكلات المفتوحة (Open-ended Tasks): في سياقات الصف الدراسي، قد يصمم المعلمون مهمات أدائية يقيسون من خلالها مهارات التفكير العلمي لدى الطلاب بشكل جماعي أو فردي. على سبيل المثال، قد يُطلب من مجموعة طلاب التخطيط لتجربة حول أفضل conditionsلنمو نبتة معينة، ثم تنفيذ التجربة فعليًا على مدى بضعة أسابيع وتسجيل نتائجهم وتقديم تقرير نهائي.

خلال هذه العملية، يقوم المعلم بملاحظة مساهمة كل طفل في صياغة الفرضيات وفي تصميم خطوات التجربة وفي كيفية تدوين البيانات وتحليلها، ثم يُقيّم أداءهم بناءً على معايير مثل: هل استطاع الطالب تحديد المشكلة بوضوح؟ هل صاغ فرضية قابلة للاختبار؟ هل كان تصميمه للتجربة منهجيًا؟ هل استندت استنتاجاته إلى البيانات بالفعل أم قفز إلى تعميمات غير مدعومة؟ ومثل هذه التقييمات العملية غالبًا ما تكون جزءًا من منهج العلوم الحديث الذي يركز على تنمية مهارات الاستقصاء العلمي بنفس قدر التركيز على المحتوى المعرفي.

ويشير Duschl et al. (2007) إلى أن توفير فرص تقييم أدائية متنوعة يضمن أن يتمكن الطالب من إظهار مهاراته العلمية في سياقات مختلفة، وليس فقط في الامتحانات التحريرية التقليدية، من خلال استخدام مزيج من هذه الأساليب المختلفة في التقييم – الاختبارات الموضوعية، الملاحظات العملية، المقابلات النوعية، والمهمات المفتوحة – يمكن للباحثين والمعلمين الحصول على صورة شاملة ومتوازنة عن مستوى مهارات التفكير العلمي لدى الطفل. فكل أداة من أدوات القياس قد تبرز جانبًا معينًا من مهارات الطفل أو تمنحنا منظورًا مختلفًا عنها. على سبيل المثال، قد يحصل طفل ما على درجة متوسطة في اختبار مكتوب للمهارات العلمية، لكن ملاحظته أثناء نشاط مختبر عملي تكشف أنه يمتلك مهارات تنفيذية جيدة ولكنه يواجه صعوبة في التعبير عنها كتابيًا.

وقد يظهر طفل آخر إتقانًا نظريًا للمفاهيم العلمية في المقابلة، لكنه يخطئ عند التطبيق العملي. لذا فإن التقييم المتنوع ضروري لضمان دقة الحكم على قدرات التفكير العلمي. كما تساعد نتائج هذه التقييمات في توجيه التدريس بحيث يركز على المهارات التي تحتاج إلى تطوير أكثر، وقياس فعالية البرامج التعليمية الموجهة لتنمية التفكير العلمي في هذه المرحلة العمرية المهمة، إن بناء أساس قوي من مهارات التفكير العلمي منذ الطفولة المبكرة يمهد الطريق أمام الطالب لمواصلة تعلم العلوم بفاعلية في المراحل اللاحقة، ويزوّده بأدوات عقلية تساعده على فهم العالم واتخاذ قرارات مبنية على التفكير النقدي المنطقي في حياته اليومية.

تطور نظم التقييم التربوي: من الأساليب التقليدية إلى النظم الذكية

الترقيم الدولي الموحد للطباعة ISSN: 3009-612X
الترقيم الدولي الموحد الالكتروني E. ISSN: 3009-6146
المجلة معرفة على قاعدة المجلات العلمية فى بنك المعرفة المصرى
وقاعدة بيانات Edu Search دار المنظومة

شهدت نظم التقييم التربوي تحولات جذرية بالتوازي مع التطورات في نظريات التعلم والتقدم التقني. في النظم التعليمية التقليدية، اعتمدت معظم أساليب التقييم على الاختبارات التحريرية القياسية كالأسئلة المقالية والاختيار من متعدد، بهدف قياس مدى تذكر المتعلم للمعلومات، هذه الأساليب التقليدية على الرغم من شيوعها ونجاحها النسبي في قياس التحصيل المعرفي الأساسي، تعرضت لانتقادات متزايدة نظرًا لعدة قصور متأصلة فيها ، فقد أشار عدد من الباحثين إلى أن التقييم التقليدي يوفر لقطات محدودة (snapshots) لأداء الطالب ولا يقدم رؤية شاملة أو عميقة عن تعلمه، كما أنه غالبًا غير متكيف مع خلفيات المتعلمين ومهاراتهم الفردية، ويعتمد صيغة موحدة قد لا تناسب احتياجات الجميع (Swiecki et al., 2022)

إضافة إلى ذلك، ترتبط الامتحانات التقليدية بثقافة التلقين والاختبار أكثر من ارتباطها بمهارات التفكير الناقد وحل المشكلات المطلوبة في عالم اليوم، ومع دخول القرن الحادي والعشرين وظهور مفهوم مهارات القرن 21، برز توجه نحو تطوير أساليب تقييم بديلة تعكس مهارات الفهم العميق والتطبيق والإبداع، برزت مثلاً التقييمات الأدائية التي تتضمن مشاريع أو تجارب مختبرية أو عروضًا شفهية لقياس مهارات التحليل والتفكير العلمي.

كذلك ازداد الاعتماد على التقويم التكويني (Formative Assessment) أثناء العملية التعليمية لتقديم تغذية راجعة فورية تدعم التعلم المستمر، ومع التقدم التقني ظهرت أنظمة التقييم المحوسبة التي سمحت بإجراء الاختبارات وتصحيحها إلكترونيًا، مما زاد من كفاءة التقييم وتوفير البيانات الفورية عن أداء الطلاب ، ثم تطور الأمر نحو الاختبارات التكيفية الحاسوبية (Computerized Adaptive Testing) التي تضبط مستوى صعوبة الأسئلة آنيًا وفق مستوى أداء المتعلم، لتحقيق قياس أدق لقدراته، وقد مهدت هذه الخطوات الطريق لظهور نظم التقييم الذكية المعتمدة على تقنيات الذكاء الاصطناعي وتعلم الآلة.

توفر هذه النظم إمكانات غير مسبوقة في تحليل استجابات الطالب وتشخيص نواحي القوة والضعف بصورة تلقائية وفورية، على سبيل المثال، يمكن لنظام تقييم ذكي تحليل أنماط إجابات الطالب في اختبار علمي لتحديد المفاهيم العلمية التي يواجه صعوبة في فهمها، وبالتالي توجيه المعلم أو المتعلم نفسه إلى تلك المفاهيم لمعالجتها. (Yoon et al., 2024).

كما تستطيع خوارزميات الذكاء الاصطناعي رصد الأخطاء الشائعة في إجابات الطلاب واقتراح تغذية راجعة تلقائية لكل متعلم بحسب احتياجاته ، وتمتاز هذه النظم الذكية أيضًا بأنها توفر صورة مستمرة عن تعلم الطالب بدلًا من الاقتصار على لقطة وحيدة في اختبار موحد، مما يسمح بقياس النمو في المهارات والفهم عبر الزمن، وقد خلُص Swiecki et al (2022) إلى أن التقييمات المعززة بالذكاء الاصطناعي بإمكانها معالجة كثير من عيوب التقييم التقليدي، مثل عبء إعداد الاختبارات وتصحيحها على المعلم، وعدم ملاءمة الاختبار الموحد

الترقيم الدولي الموحد للطباعة ISSN: 3009-612X
الترقيم الدولي الموحد الالكتروني E. ISSN: 3009-6146
المجلة معرفة على قاعدة المجلات العلمية فى بنك المعرفة المصرى
وقاعدة بيانات Edu Search دار المنظومة

لجميع المتعلمين، وتحول التقييم إلى عملية أكثر تكيفًا وإنصافًا، بناءً على ذلك يمكن القول إن نظم التقييم التربوي تطورت من نهج يقيس التعلم كناتج ثابت إلى نهج أحدث يوظف التقنية والذكاء الاصطناعي لقياس التعلم كعملية مستمرة ديناميكية، تجمع بين التشخيص والعلاج الفوري وتكيف مع حاجات المتعلم الفردية. تعلم الآلة والذكاء الاصطناعي في التعليم: التطبيقات في التشخيص والتقييم.

أحدث تطور تقنيات الذكاء الاصطناعي (AI) وتعلم الآلة (Machine Learning)نقلة نوعية في قدرات الأنظمة التعليمية على التشخيص والتقييم. تُعرَّف خوارزميات تعلم الآلة بأنها خوارزميات حاسوبية تمكن الآلة من استخلاص أنماط معرفية من البيانات واتخاذ قرارات أو تنبؤات بناءً على تلك الأنماط دون برمجة صريحة لكل حالة (Yoon et al., 2024) وقد وجدت هذه الخوارزميات تطبيقات واسعة في مجال التعليم، لا سيما في تشخيص تعلم الطلبة وتقييمه

ومن أوائل التطبيقات البارزة دمج التعلم الآلي في أنظمة الدروس الخصوصية الذكية (Intelligent Tutoring Systems) التي تقوم بتحليل تفاعل الطالب مع المادة الدراسية لحظة بلحظة، مثل إجاباته على الأسئلة أو حله للمسائل، وذلك بهدف تشخيص أخطائه أو مفاهيمه الخاطئة وتقديم إرشادات تصحيحية ملائمة ، على سبيل المثال، يمكن لخوارزمية تعلم آلة أن تتبع أخطاء الطالب في مسائل الرياضيات أو العلوم الاستقصائية، وتستنتج وجود خلل لديه في مفهوم علمي معين أو خطوة من خطوات التفكير العلمي، ثم تزوده بتغذية راجعة تصحيحية فورية حول هذا الخلل .(Kulik, 2016).

وتتميّز مثل هذه الأنظمة بقدرتها على التعلم من أداء الطالب نفسه؛ فمع كل تفاعل جديد، تحسن الخوارزمية نموذجها التشخيصي لقدرات الطالب بشكل أكثر دقة، وقد ظهرت كذلك خوارزميات متقدمة فيما يعرف بـنماذج التصنيف التشخيصي (Diagnostic Classification Models) في مجال القياس التعليمي، حيث تهدف إلى تصنيف الطلبة تبعًا لإتقانهم لمجموعة من السّمات أو المهارات الإدراكية اللازمة لحل المشكلات ، وقد بَيَّنت دراسة (Yoon et al., 2024) أن تقنيات تعلم الآلة الإشرافية يمكنها تحسين دقة تصنيف الطلاب تشخيصيًا حتى مع محدودية البيانات المتاحة للتدريب، وذلك عن طريق توليد بيانات إضافية مبنية على الأنماط المستخرجة

وهذا يعني أن الخوارزميات تستطيع تعويض ندرة عينات الأداء الطلابي من خلال نمذجة احتمالية لأداء الطالب في ظروف مختلفة، مما يزيد من موثوقية التشخيص. وإلى جانب التشخيص الفردي، أتاحت تقنيات تحليل البيانات الضخمة التعليمي المستندة إلى تعلم الآلة إمكانية رصد اتجاهات جماعية في تحصيل الطلاب؛ فمثلًا يمكن لخوارزميات التحليل العنقودي (Clustering) تجميع الطلاب حسب أنماط تعلمهم أو أخطائهم الشائعة، مما يساعد في تخصيص خطط التدريس وفق تلك الأنماط.

الترقيم الدولي الموحد للطباعة ISSN: 3009-612X
الترقيم الدولي الموحد الالكتروني E. ISSN: 3009-6146
المجلة معرفة على قاعدة المجلات العلمية فى بنك المعرفة المصرى
وقاعدة بيانات Edu Search دار المنظومة

أما في مجال التقييم، فقد استُخدمت تقنيات معالجة اللغة الطبيعية (NLP) لتقييم الإجابات المفتوحة أو المقالية للطلاب من خلال فهم المحتوى النصي الذي يكتبه الطالب ومقارنته بمعايير الأداء المطلوبة، مما يوفر درجة تقييمية آلية وتغذية راجعة نوعية. إضافة إلى ذلك، ظهرت أنظمة تقييم قادرة على رصد تفكير الطالب أثناء حل المشكلات من خلال تتبع خطواته أو تفاعلاته في بيئة افتراضية، فيما يعرف أحيانًا بـ "التقييم الخفي (Stealth Assessment)" حيث يجمع النظام بيانات عن أداء الطالب دون أن يشعر بأنه يخضع لاختبار مباشر.

هذه البيانات الضخمة الملتقطة يمكن لخوارزميات التعلم العميق تحليلها لاستخلاص دلائل على مستوى إتقان الطالب لمهارات عليا كالتفكير التحليلي أو التفكير المنظومي، وقد أكدت مراجعات منهجية حديثة أن تطبيقات الذكاء الاصطناعي في التعليم تتركز بشكل كبير في مجال التقييم والتغذية الراجعة الشخصية، الأمر الذي يعكس أهمية هذه التقنيات في تقديم تقييم مستمر ودقيق يمكن توظيفه لتحسين مخرجات التعلم (Moroianu et al., 2023)

على سبيل المثال، أشارت دراسة كل من Moroianu et al,. (2023) إلى أن التقييم والتقويم هو أكثر المجالات التي شهدت توظيفًا للذكاء الاصطناعي في التعليم العالي، حيث سُخِّرت الخوارزميات لتوفير تغذية راجعة تلقائية ومخصصة للطلاب، فضلًا عن التنبؤ بمخاطر تعثرهم الأكاديمي

كما أسهم توظيف الذكاء الاصطناعي في التقييم في جعل التفكير غير المرئي مرئيًا؛ فكما يصف أحد خبراء التربية العلمية، تساعد أدوات الذكاء الاصطناعي في جعل عملية تفكير الطلاب أكثر وضوحًا للمعلم عبر تتبع استدلالاتهم خطوة بخطوة مما يمكن المعلم من فهم طريقة تفكير الطالب وتصميم تدخلات تعليمية أكثر فعالية ، فتعلم الآلة والذكاء الاصطناعي يمثلان أدوات ثورية لتعزيز قدرات النظم التعليمية في تشخيص تعلم الطلاب وتقويم مهاراتهم العليا، شريطة تصميم هذه الأدوات واستخدامها بشكل منهجي وأخلاقي يراعي خصوصية البيانات ويضمن العدالة والشفافية في التقييم (Bennett & Royce, 2025) .

العلاقة بين التفكير العلمي والتشخيص ونظم التقييم الذكية

هناك ترابط وثيق بين كل من مهارات التفكير العلمي وعملية التشخيص ونظم التقييم الذكية حيث يؤثر كل منها في الآخر بصورة تبادلية ضمن سياق تعليم العلوم الحديث. فمن ناحية، تنمية مهارات التفكير العلمي لدى المتعلمين تعتمد بشكل كبير على القدرة على تشخيص مستوى امتلاكهم لهذه المهارات وتشخيص صعوبات التعلم المتعلقة بها. دون تشخيص دقيق، يبقى تعزيز التفكير العلمي عملية عشوائية تفتقر إلى التركيز على احتياجات الطالب المحددة.

وهنا تبرز أهمية التشخيص التربوي كعملية مستمرة يتم عبرها جمع معلومات تفصيلية حول تفكير الطالب العلمي، بما في ذلك قدرته على الملاحظة وطرح الأسئلة وفرض الفرضيات والتفسير المنطقي للأدلة .التقييمات الذكية المعتمدة على الذكاء الاصطناعي جاءت لتحدث نقلة نوعية في إمكانيات هذا

الترقيم الدولي الموحد للطباعة ISSN: 3009-612X
الترقيم الدولي الموحد الالكتروني E. ISSN: 3009-6146
المجلة معرفة على قاعدة المجلات العلمية فى بنك المعرفة المصرى
وقاعدة بيانات Edu Search دار المنظومة

التشخيص؛ إذ تسمح خوارزميات تعلم الآلة المدمجة في منصات التقييم بجمع وتحليل كم هائل من بيانات أداء الطالب بشكل فوري، مما يوفر صورة دقيقة عن أنماط تفكيره العلمي .على سبيل المثال، إذا كان أحد الطلاب يواجه صعوبة في صياغة فرضيات قابلة للاختبار أثناء الأنشطة الاستقصائية، يمكن لنظام تقييم ذكي رصد هذه الصعوبة من خلال تحليله لأسئلة الطالب أو توقعاته أثناء التجربة. يقوم النظام حينها بتشخيص المهارة المحددة التي يحتاج الطالب إلى دعم فيها (في هذه الحالة مهارة صياغة الفرضيات) وإعلام المعلم أو اقتراح نشاط علاجي مناسب، وبالتالي فإن نظم التقييم الذكية تعمل كأداة تشخيصية تكشف بدقة عن نقاط القوة ونقاط الضعف في تفكير الطالب العلمي .(Jones, 2021)

ومن ناحية أخرى، يتطلب تصميم نظم تقييم ذكية فعّالة فهمًا عميقًا لطبيعة التفكير العلمي ذاته، أي تحديد ما الذي سنقوم بتشخيصه وتقويمه. لذا فالعلاقة بين التفكير العلمي والتقييم الذكي علاقة تصميمية أيضًا؛ حيث يجب تضمين معايير وأبعاد التفكير العلمي (مثل مهارات الاستقصاء والاستدلال والتفكير الناقد) ضمن نماذج التقييم الذكي لضمان أن ما يقيسه النظام يتوافق مع جوهر التفكير العلمي المنشود.

في السياق نفسه، يضفي التشخيص التربوي بعدًا زمنيًا على العلاقة بين التفكير العلمي والتقييم الذكي. فالتشخيص ليس حدثًا منفصلًا عن التقييم، بل هو وظيفة من وظائف عملية التقييم حين يكون التقييم مستمرًا وتكوينيًّا. فالتشخيص يزوّد نظام التقييم الذكي ببيانات راجعة تمكنه من تعديل مسار التعليم لكل طالب بشكل فردي (التعلم الشخصي). وبالمقابل، يوفر نظام التقييم الذكي بيئة للتشخيص الفوري والمتواصل تجعل المعلم على دراية بتطور تفكير الطالب العلمي خطوة بخطوة.

وهنا تتجلى العلاقة التفاعلية " تحسين التفكير العلمي هدف، والتشخيص وسيلة، والتقييم الذكي أداة" لتحقيق تلك الوسيلة بكفاءة وموثوقية عالية.

وخلصت دراسة Shute & Ventura (2013) إلى أن الاستفادة القصوى من نظم التقييم الذكية تتحقق عندما تُصمم بغرض تشخيص عمليات التفكير وليس فقط نتائج الإجابات الصحيحة أو الخاطئة) أي أن النظام الذكي لا يكتفي بتحديد خطأ الطالب، بل يحلل نوع الخطأ وسببه المرتبط بطريقة تفكير الطالب، سواء كان خطأ ناتجًا عن افتراض غير صحيح أو عن استدلال ناقص. بذلك يصبح التقييم عملية تعلم في حد ذاتها، حيث يتعرف الطالب على أسلوب تفكيره ويتأمل فيه بفضل التشخيص الفوري من النظام.

فالتفكير العلمي يوجه تصميم التقييم الذكي من حيث ما ينبغي قياسه وكيف، والتقييم الذكي يقدم بيانات التشخيص، والذي بدوره يستخدم لتحسين التفكير العلمي وتعزيزه لدى المتعلم، هذه الحلقة المتكاملة تمثل إطارًا نظريًا حديثًا لتحسين تعلم العلوم، تستفيد فيه التربية من إمكانات التقنية الحديثة لتحقيق تعليم أكثر فعالية وملاءمة لاحتياجات الطلاب الفردية.

الترقيم الدولي الموحد للطباعة ISSN: 3009-612X
الترقيم الدولي الموحد الالكتروني E. ISSN: 3009-6146
المجلة معرفة على قاعدة المجلات العلمية فى بنك المعرفة المصرى
وقاعدة بيانات Edu Search دار المنظومة

تعليم العلوم في المرحلة الابتدائية في مصر: الواقع والتحديات

يمثل تحسين تعليم العلوم في المرحلة الابتدائية في مصر تحديًا استراتيجيًا في ظل السعي لإعداد جيل قادر على التفكير العلمي والإبداع، أشارت الأدبيات التربوية إلى وجود عدد من التحديات المزمنة التي تواجه تعليم العلوم وتقويم مهارات التفكير العليا في السياق المصري ، من أبرز هذه التحديات تركيز المناهج والممارسات التقليدية على الحفظ والتلقين لغرض الامتحان على حساب تنمية مهارات التفكير الناقد والاستقصاء، فقد أشار تقرير لمركز(Loveluck, 2012) ، إلى أن أساليب التدريس السائدة في النظام التعليمي المصري ظلت لعقود توجه الطلاب نحو الاستظهار والاستجابة الامتحانية بدلًا من تشجيعهم على التفاعل النقدي مع المحتوى

ونتج عن ذلك ضعف في اكتساب المتعلمين لمهارات التفكير العلمي؛ إذ نادرًا ما يتاح لتلاميذ المرحلة الابتدائية ممارسة التجريب الحر أو طرح الأسئلة أو مناقشة الأفكار العلمية بصورة مفتوحة، ويُضاف إلى ذلك تحدٍ آخر يتمثل في قصور نظم التقييم الحالية في قياس مهارات التفكير العليا، فالتقويم في المرحلة الابتدائية بمصر ما زال يعتمد بشكل كبير على اختبارات تحصيلية معيارية تركز على استرجاع المعلومات المباشرة، دون أن تقيس قدرة التلميذ على التحليل أو التطبيق أو الاستنتاج .أضف إلى ذلك محدودية استخدام التقييم التكويني أو التشخيصي بشكل منهجي؛ فالمعلم غالبًا لا يمتلك أدوات فعالة تمكنه من رصد تطور تفكير الطالب بشكل مستمر خلال الفصل

وتؤكد دراسة استكشافية حديثة حول مدارس المتفوقين STEM في مصر (El Nagdi & Roehrig, 2022)هذا الخلل، إذ كشفت عن وجود فجوة بين سياسات التعليم الداعية إلى التعلم القائم على المشروعات والاستقصاء، وبين واقع التقييم الذي يظل تقليديًا معتمدًا على اختبارات غير متماشية مع تلك الأساليب الحديثة

وإن كان هذا السياق خاصًا بمدارس STEM، إلا أنه يعكس عمومًا ظاهرة عدم مواءمة التقييم مع أهداف تنمية المهارات العليا حتى في بيئات يفترض أنها متقدمة، ومن التحديات الأخرى في الواقع المصري أيضًا ارتفاع كثافات الفصول وضعف البنية التحتية المعملية في كثير من المدارس، مما يصعّب تنفيذ أنشطة استقصاء علمي عملية.

كما أن تدريب المعلمين على توجيه التفكير العلمي لدى التلاميذ ما يزال غير كافٍ، حيث لم تتضمن برامج إعداد المعلم بشكل عميق استراتيجيات تعليم مهارات التفكير ولا طرق التقييم الحديثة لهذه المهارات، علاوة على ذلك، يواجه إدخال التقنيات التعليمية الذكية في المدارس صعوبات تتعلق بالكلفة والتجهيز وضعف الاتصال بالإنترنت في بعض المناطق، مما يعني أن تطبيق نظم التقييم الذكية ليس بالأمر اليسير حاليًا على نطاق واسع.

الترقيم الدولي الموحد للطباعة ISSN: 3009-612X
الترقيم الدولي الموحد الالكتروني E. ISSN: 3009-6146
المجلة معرفة على قاعدة المجلات العلمية فى بنك المعرفة المصرى
وقاعدة بيانات Edu Search دار المنظومة

ومع ذلك، هناك بوادر إيجابية متمثلة في تبني وزارة التربية والتعليم مؤخرًا إطارًا جديدًا للمناهج يركز على نواتج تعلّم ترتبط بالمهارات وليس المحتوى فقط، مثل مهارات التفكير الناقد والتعاون، حيث أدخلت مناهج العلوم للمرحلة الابتدائية أنشطة مبسطة في الاستقصاء والمشروعات الصغيرة التي تتحدى التلميذ للتفكير وحل المشكلات، لكن تبقى آلية التقويم غير متسقة كليًا مع هذه المناهج المطورة، حيث لا تزال الامتحانات الموحدة هي المعيار لتقييم الطلاب في نهاية السنة الدراسية ، هذا التناقض يخلق ثغرة بين ما يُعلَّم وما يُقوَّم، وقد يثبط من جهود المعلمين والطلاب في التركيز على المهارات إذا كانت الامتحانات لا تكافئ التفكير العلمي بشكل مناسب.

أيضًا يبرز تحدي الثقافة المدرسية وأولياء الأمور الذين اعتادوا على أسلوب التعليم التقليدي، مما يجعلهم أقل تقبلًا لممارسات تعليمية جديدة كتلك القائمة على المشروعات أو لنظم تقييم غير مألوفة. ففي كثير من الأحيان، ينظر أولياء الأمور إلى الاختبارات الرقمية أو التقييم التكويني المستمر بعين الشك، مفضلين الاختبار الورقي التقليدي ظنًا بأنه أكثر موثوقية.

في ضوء هذه التحديات، يتضح أن الواقع المصري يتطلب معالجة شاملة تشمل تطوير أساليب التدريس والتقويم معًا، ونظم التقييم الذكية يمكن أن تكون جزءًا من الحل إذا ما توافرت لها البيئة الداعمة، إذ ستساعد المعلمين في تشخيص تعلم الطلاب على نحو فردي حتى في الفصول الكبيرة، كما ستوفر بيانات موضوعية عن مهارات التفكير العلمي لكل تلميذ يمكن استخدامها لتحسين التدريس، ولابد أن يسبق ذلك تهيئة المدارس والمعلمين وتطوير البنية التحتية التقنية.

ومما سبق يمكننا القول إن تعليم العلوم في المرحلة الابتدائية بمصر يمر بمرحلة تحول تسعى للتخلي عن إرث التلقين نحو تعليم قائم على تنمية المهارات والفهم، غير أن هذا التحول يواجه عقبات تتعلق بالتقييم بالدرجة الأولى، ومن هنا تنبع حاجة البحث الحالي إلى إيجاد أساليب وأدوات جديدة – كالنظم الذكية – للمساعدة في تقويم مهارات التفكير العلمي وتشخيصها في ظل الواقع المحلي، بما يدعم جهود تطوير تعليم العلوم بشكل متكامل

مصطلحات البحث:

أ- مهارات التفكير العلمي لدى تلاميذ المرحلة الابتدائية :

تعرف الباحثة مهارات التفكير العلمي إجرائياً في البحث الحالي على أنها "مجموعة الأداءات العقلية والسلوكية التي يُظهرها تلميذ الصف السادس الابتدائي عند تعامله مع مشكلات او مواقف علمية من خلال تطبيق منهجي لثماني مهارات رئيسية، وهي: (الملاحظة الدقيقة – صياغة الفرضيات – تصميم التجارب – التنبؤ – تفسير البيانات – الاستنتاج – ضبط المتغيرات – التواصل العلمي)، ويُقاس تمكن الطفل من هذه المهارات من خلال:

الترقيم الدولي الموحد للطباعة ISSN: 3009-612X
الترقيم الدولي الموحد الالكتروني E. ISSN: 3009-6146
المجلة معرفة على قاعدة المجلات العلمية فى بنك المعرفة المصرى
وقاعدة بيانات Edu Search دار المنظومة

- أدائه لمهام وأنشطة صفية أو عملية تتضمن خطوات البحث العلمي.
- قدرته على جمع بيانات موثوقة باستخدام الحواس (الملاحظة).
- إنتاج فرضيات مبدئية قابلة للاختبار.
- تصميم تجارب تتسم بالمنهجية والتحكم في المتغيرات.
- التنبؤ بنتائج منطقية استناداً إلى معطيات سابقة.
- تحليل البيانات واستنتاج دلالات عملية منها.
- إصداء استنتاجات مدعومة بالأدلة.
- نقل المعارف العلمية شفهياً وكتابياً برسوم أو تقارير.

ب- خوارزميات تعلّم الآلة:(Machine Learning Algorithms)

أساليب حسابية تمكّن الأنظمة من التعلّم تلقائيًا من البيانات وتحسين أدائها تدريجيًا دون برمجة صريحة لكل خطوة (Nguyen & Walker, 2016; Baker & Yacef, 2009)، وتُستخدم في البحث الحالي لتصنيف التلاميذ إلى مستويات التفكير العلمي (متفوّق – متوسط – ضعيف) استنادًا إلى أدائهم في أدوات القياس المستخدمة في البحث، وقد تم توظيف خوارزمية الغابة العشوائية (Random Forest)لكفاءتها العالية في التعامل مع بيانات متعددة الأبعاد، وقدرتها على التقاط العلاقات غير الخطّية وإظهار أهمية الميزات.

ت- التقييم الذكي :(Intelligent Assessment) :

تُعرفه الباحثة إجرائياً في البحث الحالي على أنه "منظومة رقمية تشخيصية تعتمد خوارزميات تعلّم الآلة لتحليل درجات التلاميذ في الأدوات الثلاث (اختبار تحصيلي، مهام أدائية، بطاقة ملاحظة)، ودمج درجات المهارات الثماني، ثم توليد تصنيف نهائي يعكس بدقّة أداء الطالب في التفكير العلمي. يقوم النموذج بتحويل المؤشرات السلوكية والأدائية إلى قرارات تصنيفية مدعومة بدليل إحصائي".

فروض البحث

1- تتحقق لبطارية قياس مهارات التفكير العلمي (الاختبار التحصيلي – المهام الأدائية – بطاقة الملاحظة) مؤشراتٌ سيكومترية مقبولة ودالّة إحصائيًا

الترقيم الدولي الموحد للطباعة ISSN: 3009-612X
الترقيم الدولي الموحد الالكتروني E. ISSN: 3009-6146
المجلة معرفة على قاعدة المجلات العلمية فى بنك المعرفة المصرى
وقاعدة بيانات Edu Search دار المنظومة

2- يتمتع نموذج تعلم الاله المستخدم (Random Forest) بقدرة تصنيفية دالة إحصائيًا في تحديد مستوى التفكير العلمي لدى التلاميذ (ضعيف – متوسط-ممتاز) بنسبة دقة لا تقل عن 80%.

3- توجد علاقة ارتباطيه موجبة ودالة إحصائيًا بين متوسط درجات التلاميذ في مهارات التفكير العلمي والتصنيف الناتج عن نموذج تعلم الاله.

4- توجد فروق ذات دلالة إحصائية بين متوسط درجات التلاميذ في المهارات العلمية تبعاً لمستوى تصنيفهم النهائي (ضعيف – متوسط – ممتاز).

5- تسهم بعض المهارات الفرعية للتفكير العلمي بدرجة أكبر من غيرها في التنبؤ بمستوى التفكير العلمي لدى التلاميذ.

عينة البحث:

تتكون عينة البحث الميدانية من (400) تلميذ من تلاميذ الصف السادس بالمرحلة الابتدائية الملتحقين بمدارس التعليم الابتدائي في محافظة البحيرة بجمهورية مصر العربية، يمثل هؤلاء التلاميذ الفئة المستهدفة للدراسة نظرًا لملاءمة المرحلة العمرية لمهام القياس المستخدمة (الاختبار التحصيلي، المهام الأدائية، بطاقة الملاحظة) ولارتباطها بتنمية مهارات التفكير العلمي في نهاية الحلقة الثانية من التعليم الأساسي، ويوضح جدول (3) توزيع العينة طبقاً للنوع

جدول (3) توزيع العينة الأساسية للدراسة طبقاً لمتغير النوع

النوع	العدد	النسبة المئوية
البنات	206	51.5%
البنون	194	48.5%
العدد الكلى	400	100%

يُظهر التوزيع النوعي توازنًا مقبولًا بين الذكور والإناث، بما يدعم استقرار تقديرات النموذج لاحقًا.

وتتسم العينة بتجانس عمري مناسب حيث ينتمي أفراد العينة لصفٍ دراسي واحد؛ إذ بلغ متوسط العمر الزمني 11.62 سنة، بانحراف معياري 0.49 سنة، ويعكس ذلك نطاقًا عمريًا ضيقًا ومتوقعًا لتلاميذ الصف السادس، بما يقلل من أثر التغاير العمري على مخرجات القياس والتحليل.

معايير الإدراج والاستبعاد

- الإدراج :تلاميذ الصف السادس المنتظمون حضورًا خلال فترة تطبيق أدوات القياس، والقادرون على إكمال الاختبار التحصيلي، والمهام الأدائية المعيارية، والخضوع لبطاقة الملاحظة وفق بروتوكول موحّد.

الترقيم الدولي الموحد للطباعة ISSN: 3009-612X
الترقيم الدولي الموحد الالكتروني E. ISSN: 3009-6146
المجلة معرفة على قاعدة المجلات العلمية فى بنك المعرفة المصرى
وقاعدة بيانات Edu Search دار المنظومة

- الاستبعاد :الحالات التي تغيب أصحابها عن أحد تطبيقات القياس الأساسية، أو لم تستكمل بياناتهم بما يحول دون حساب الدرجات الموحدة أو المؤشر المركب.

وللتحقق من الخصائص السيكومترية لأدوات القياس، تم استخدام عينة لحساب الخصائص السيكومترية قوامها (ن = 40) تلميذ وتلميذه من تلاميذ الصف السادس الابتدائي بمحافظة البحيرة، وبنفس مواصفات العينة الأساسية من حيث المرحلة الدراسية والسياق المدرسي وإجراءات التطبيق، هدفت هذه العينة إلى فحص ملاءمة الفقرات وسلامة الصياغة، والتأكد من ثبات الأدوات وصدقها قبل التطبيق على العينة الأساسية.

منهج البحث:

اعتمد البحث على المنهج الوصفي التحليلي، مع توظيف خوارزميات تعلم الآلة ضمن إطار تحليل البيانات التربوية.Educational Data Mining

خطوات البحث

شملت خطوات البحث ما يلي:

- جمع البيانات :تم إعداد بطارية أدوات تقيس مهارات التفكير العلمي شملت اختباراً تحصيلياً مقنّناً، ومهاماً أدائية عملية، وبطاقة ملاحظة للمهارات الصفية أثناء التجربة العملية. وجُمعت بيانات كل تلميذ عبر هذه الأدوات.
- تحرير الدرجات :بعد ذلك جُمّعت درجات كل أداة على حدة، وحُسبت المتوسطات الموحّدة لكل من ثماني مهارات للتفكير العلمي (مثل صياغة الفرضية، التصميم التجريبي، الخ) ومؤشر المجموع المركب.
- تحليل البيانات الوصفية :أُجري تحليل وصفي للبيانات (متوسطات، انحراف معياري) لفهم توزيع درجات المهارات ومستوى التلاميذ.
- بناء النموذج التدريبي :استُخدمت خوارزمية Random Forestلبناء نموذج تصنيف التلاميذ إلى ثلاث فئات (ضعيف–متوسط–ممتاز). قُسِّمَت البيانات إلى مجموعات تدريب واختبار (مثلاً 80% تدريب، 20% اختبار)، واستخدمت مكتبة scikit-learn لتنفيذ الغابة العشوائية.
- تقييم النموذج :وُفّرت مقاييس الأداء الأساسية للنموذج، مثل الدقة(accuracy) ، والاستدعاء(recall) ، ودقة التنبؤ(precision) ، ومعامل F1 ، إضافةً إلى مصفوفة الالتباس لقياس مدى صلاحية التصنيف.
- تحليل النتائج :نُفّذت اختبارات إحصائية لتحليل الفرضيات (مثل اختبار الارتباط بين الدرجات والتصنيف، واختبار ANOVA لفروق المتوسطات)، بالإضافة إلى حساب أهمية المتغيرات (feature importance)من نموذج الغابة العشوائية لتحديد مدى مساهمة كل مهارة فرعية في التنبؤ بمدى التفكير العلمي.

الترقيم الدولي الموحد للطباعة ISSN: 3009-612X
الترقيم الدولي الموحد الالكتروني E. ISSN: 3009-6146
المجلة معرفة على قاعدة المجلات العلمية فى بنك المعرفة المصرى
وقاعدة بيانات Edu Search دار المنظومة

واعتمدت الباحثة على خوارزمية الغابة العشوائية(Random Forest) ، وهي خوارزمية تعلم آلي تجمع عدة أشجار قرار (Decision Trees) في نموذج واحد .فهي تنشئ «غابة» تتكون من مئات الأشجار البسيطة، يعمل كل منها على جزء مختلف من البيانات، ثم يجمع نتائجها النهائيّة بتصويت الأغلبية لضمان تنبؤ أكثر دقة، من أهم مميزات الغابة العشوائية أنها مرنة وسهلة الاستخدام وتحقق دقة عالية في التصنيف حتى دون ضبط دقيق للمعاملات (hyperparameters) كما يمكنها التعامل مع المهام التصنيفية والانحدار على حد سواء، وتقلل خطر الإفراط في التخصيص (overfitting) بفضل اعتمادها على تصويت مجموعة من الأشجار مستقلة نسبياً.

في سياق قياس مهارات التفكير العلمي، يبرر اختيار الغابة العشوائية أنها تلائم البيانات المتعددة الأبعاد (دمج بيانات الاختبار والتحصيل العملي والملاحظات) وتوفر فهماً لكيفية مساهمة كل مهارة فرعية في التصنيف النهائي، فعلى سبيل المثال، تتيح خوارزمية الغابة العشوائية حساب أهمية الميزات (مقدار تأثير كل متغير) تلقائياً بعد التدريب، مما يساعد الباحث على تفسير النموذج ومعرفة المهارات الأكثر تأثيراً في تصنيف مستوى الطلاب ، وهذه الخاصية مهمة في البحوث التربوية لأنها تقدم رؤية عملية حول أي المهارات تسهم أكثر في تفكير الطالب العلمي (Donges,2024)

تصميم أداة متكاملة لقياس مهارات التفكير العلمي لدى تلاميذ المرحلة الابتدائية (الصف السادس الابتدائي)

مبررات تصميم الأداة

يعد تنمية مهارات التفكير العلمي لدى التلاميذ في مرحلة التعليم الأساسي هدفًا أساسيًا في مناهج العلوم الحديثة، لما لهذه المهارات من دور محوري في إعداد متعلم قادر على حل المشكلات واتخاذ القرارات المبنية على الأدلة، وتظهر نتائج العديد من الدراسات علاقة قوية بين التفكير العلمي وتحسن التحصيل الدراسي للطلبة وتؤكد ضرورة البدء في تنمية هذه المهارات منذ المراحل المبكرة

على سبيل المثال، يذكر عبد الحميد زيتون (2008) أن إكساب المتعلمين مهارات التفكير العلمي هو أحد الأهداف الرئيسة لتدريس العلوم الحديثة (عبد الحميد زيتون، 2008)، كما تُبيّن دراسة كل من (Ergin & Aktamis, 2008) أن امتلاك الطلبة لمهارات التفكير العلمي يمكّنهم من تحديد مشكلاتهم وملاحظة الظواهر وتحليلها وافتراض الفروض المناسبة واستخلاص النتائج وتفسيرها وتعميمها على مواقف جديدة، مما يسهم في إعدادهم لحل المشكلات الحياتية بطرق منهجية فعّالة.

ويتطلب تقييم مهارات التفكير العلمي أدوات تقييم متنوعة تتجاوز الاختبارات التقليدية، فالتفكير العلمي ينطوي على أداءات وسلوكيات قد لا تُقاس بدقة عبر أسئلة الاختيار المتعدد فقط

الترقيم الدولي الموحد للطباعة ISSN: 3009-612X
الترقيم الدولي الموحد الالكتروني E. ISSN: 3009-6146
المجلة معرفة على قاعدة المجلات العلمية فى بنك المعرفة المصرى
وقاعدة بيانات Edu Search دار المنظومة

لذلك جاءت فكرة تصميم أداة متكاملة (بطارية من الأدوات) تجمع بين اختبار ورقي ومهام أدائية وبطاقات ملاحظة، لضمان قياس شامل لجميع جوانب التفكير العلمي ، فاستخدام طرق تقييم متعددة يوفر صورة أكثر صدقًا لقدرات التلميذ، حيث تكمل المعلومات المستقاة من كل طريقة بعضها البعض

فالاختبارات الورقية تركز على الكشف عن قدرة التلميذ على استنتاج المعلومات من نص مكتوب، بينما تُظهر المهمة الأدائية مهارته في التجريب العملي وضبط المتغيرات، وتتيح بطاقة الملاحظة توثيق سلوكه العلمي أثناء تنفيذ النشاط ، إن هذا التكامل في أساليب التقييم يتماشى مع توصيات التربويين بالحاجة إلى تنويع طرائق التقويم لتحسين دقته وشموليته.(Hopkins et al., 1990)

ولتحديد المهارات الفرعية للتفكير العلمي المستهدفة في الدراسة الحالية ، تم الرجوع إلى الأدبيات التربوية والدراسات السابقة التي تناولت عمليات العلم ومهارات التفكير العلمي لدى المتعلمين ، أظهرت هذه الدراسات اتفاقًا واسعًا حول مجموعة من المهارات الأساسية، رغم اختلاف المسميات أحيانًا، على سبيل المثال، قدّمت فتحية اللولو (1997) قائمة بمهارات التفكير العلمي تضمنت: الملاحظة، التصنيف، الاستنتاج، التنبؤ، القياس، صياغة الفرضيات، التجريب، وعزل المتغيرات (فتحية اللولو، 1997) ، وفي السياق ذاته طوّر Padilla (1990)تصنيفًا عالميًا لمهارات عمليات العلم صنفها إلى مهارات أساسية (مثل: الملاحظة، التواصل، التصنيف، القياس، التنبؤ، الاستنتاج) ومهارات مركبة أو تكاملية (مثل: ضبط المتغيرات، الفرضيات، تفسير البيانات، التجريب)

كما أكدت دراسة رياض بهجات (2003) على أهمية مهارات مثل تحديد المتغيرات والاتصال العلمي والتفسير كجزء من التفكير العلمي لدى طلبة المرحلة المتوسطة (رياض بهجات، 2003)، وبناءً على مراجعة تلك الدراسة وكذلك دراسة كل من (إلهام الشلبي وشذى الخليفة، 2017) ؛ (Aktamis & Ergin, 2008)، تم اعتماد مجموعة من المهارات الأساسية للتفكير العلمي كأبعاد رئيسة للأداة: الملاحظة، وصياغة الفرضيات، وتصميم التجربة (التجريب)، والتنبؤ، وتفسير البيانات، والاستنتاج، وضبط المتغيرات، والتواصل العلمي، هذه المهارات تمثل مكونات عملية التفكير العلمي التي يُتوقع من تلاميذ الصف السادس إجادتها عند مشاركتهم في أنشطة العلوم، وفيما يلي تعريف دقيق لكل مهارة كما ورد في الأدبيات: (فتحية اللولو، 1997).

1- الملاحظة: (Observation)

"تعني الانتباه الحسي الدقيق والمنظّم للظواهر والأحداث بهدف جمع المعلومات والبيانات حولها "، وتتضمن الملاحظة استخدام الحواس المختلفة لوصف الخصائص أو التغيرات في الأشياء أو الأحداث. مثلًا: ملاحظة لون محلول كيميائي أو تغير درجة حرارة جسم ما. تُعد الملاحظة نقطة الانطلاق لأي بحث علمي ناجح.

الترقيم الدولي الموحد للطباعة ISSN: 3009-612X
الترقيم الدولي الموحد الالكتروني E. ISSN: 3009-6146
المجلة معرفة على قاعدة المجلات العلمية فى بنك المعرفة المصرى
وقاعدة بيانات Edu Search دار المنظومة

2- صياغة الفرضيات (Formulating Hypotheses):

"تشير إلى قدرة المتعلم على اقتراح تفسيرات مؤقتة أو حلول محتملة لعلاقة بين متغيرات"، والفرضية العلمية عبارة عن تخمين ذكي قابل للاختبار يربط بين متغير مستقل ومتغير تابع كإجابة محتملة لسؤال أو مشكلة، على سبيل المثال: "إذا زادت كمية ضوء الشمس، فإن معدل نمو النبتة سيزداد" هي فرضية تربط بين الضوء ونمو النبات.

3- تصميم التجربة/التجريب (Experimentation):

"يقصد بها قدرة التلميذ على تصميم وإجراء تجربة علمية للتحقق من صحة فرضية ما"، ويشمل ذلك تحديد خطوات التجربة والإجراءات، واستخدام المواد والأدوات بشكل صحيح، وجمع البيانات بطريقة منهجية، ويتضمن التصميم التجريبي أيضًا التحكم في المتغيرات لضمان سلامة النتائج.

4- التنبؤ (Prediction):

"هو توقع يعتمد على المعلومات والأنماط المتوفرة للتنبؤ بما قد يحدث مستقبلاً في ظروف معينة "، ويستخدم التلميذ معرفته السابقة وبيانات الملاحظة للاستدلال على نتيجة لم تُلاحظ بعد. مثال: التنبؤ بنتيجة تجربة قبل إجرائها بناءً على معطيات التجارب المشابهة السابقة.

5- تفسير البيانات (Data Interpretation):

"تعني تحليل المعطيات المستخلصة من الملاحظات أو التجارب وتنظيمها واستخلاص مدلولاتها"، تشمل هذه المهارة قراءة الجداول والرسوم البيانية، وربط النتائج بالفرضيات المطروحة، وتقديم تفسير علمي للعلاقات أو الأنماط التي تظهرها البيانات. ويمثل التفسير حلقة وصل مهمة بين ما تم ملاحظته وما يمكن استنتاجه من نتائج التجربة.

6- الاستنتاج (Conclusion/Inferring):

"هو القدرة على استخلاص نتائج عامة أو استدلالات من معلومات وأدلّة محددة "، ويستند الاستنتاج إلى الملاحظات والتفسيرات السابقة لتكوين فهم أعمق للظاهرة محل الدراسة. مثلًا: بعد إجراء تجربة حول تأثير نوع السماد على نمو النبات، يتم استنتاج أي أنواع السماد كان الأكثر فعالية ولماذا. الاستنتاج مرتبط بشكل وثيق بالتفسير، إذ يُشكل المنتج النهائي لعملية التفكير بعد جمع وتحليل البيانات.

7- ضبط المتغيرات (Controlling Variables):

"هي مهارة تحديد العوامل التي يمكن أن تؤثر في نتيجة التجربة والتحكم بها"، ويتضمن ذلك التمييز بين المتغيرات المستقلة (التي نغيرها عمداً) والمتغيرات التابعة (التي نقيس أثر التغير فيها) والمتغيرات الضابطة (الثابتة التي لا نغيرها). تمكن هذه المهارة التلميذ من تصميم تجارب منضبطة بحيث يكون أي تغيير في النتائج ناتجًا عن المتغير المستقل فقط

الترقيم الدولي الموحد للطباعة ISSN: 3009-612X
الترقيم الدولي الموحد الالكتروني E. ISSN: 3009-6146
المجلة معرفة على قاعدة المجلات العلمية فى بنك المعرفة المصرى
وقاعدة بيانات Edu Search دار المنظومة

8- التواصل العلمي :(Scientific Communication)

"يقصد به قدرة التلميذ على نقل الأفكار والمعلومات العلمية للآخرين بوضوح ودقة، سواء شفهيًا أم كتابيًا أم من خلال رسوم وأشكال توضيحية"، وتشمل هذه المهارة وصف الإجراءات والنتائج بلغته الخاصة أو باستخدام جداول ورسوم بيانية، وكذلك مناقشة النتائج مع الزملاء والمعلمين. فمثلاً: كتابة تقرير مختصر عن تجربة قام بها الطالب، أو تقديم عرض شفهي يشرح فيه ما توصّل إليه من نتائج، يعدان شكلين من أشكال التواصل العلمي. هذه المهارة مهمة لضمان أن فهم التلميذ العلمي يمكن مشاركته وتقييمه من قبل الآخرين.

هذه المهارات الثمانية تم اعتمادها كأبعاد رئيسية للأداة المتكاملة. وقد تم اختيارها استناداً إلى أسس علمية؛ فقد وجد أنها شاملة لمعظم ما أورده الباحثون في تحديد مكونات التفكير العلمي لدى المتعلمين (فتحية اللولو، 1997؛ Padilla, 1990؛ رياض بهجات، 2003)، وحرص التصميم على تغطية كل منها لضمان تقييم متوازن لقدرات التلاميذ العلمية.

وتتكون الأداة من ثلاث مكونات فرعية مكملة لبعضها، هي: اختبار ورقي تحصيلي، ومهام أدائية عملية، وبطاقة ملاحظة، صُممت هذه المكونات لتُستخدم بشكل فردي مع التلميذ، بحيث يؤدي الاختبار ويقوم بالمهام تحت إشراف المقيم الذي يستخدم أيضًا بطاقة الملاحظة أثناء الأداء، فيما يلي وصف تفصيلي لشكل كل مكون مع أمثلة توضيحية:

1 . الاختبار التحصيلي

الاختبار التحصيلي هنا هو اختبار ورقي مكتوب يهدف إلى قياس جوانب التفكير العلمي التي يمكن الكشف عنها عبر الأسئلة التحريرية، تم إعداد الاختبار ليشمل أنماطًا متنوعة من الأسئلة (اختيار من متعدد، وأسئلة قصيرة مفتوحة، وأسئلة موقفيه) تغطي كل مهارة من المهارات المستهدفة. يبلغ عدد فقرات الاختبار (40 فقرة) موزعة بالتساوي على المهارات الثماني (5 فقرات لكل مهارة)، وتُعطى كل فقرة درجة واحدة، وقد تم التأكد من ملاءمة محتوى الأسئلة لمستوى الصف السادس الابتدائي.

ويصحح الاختبار اعتمادًا على مفاتيح الإجابات النموذجية لكل سؤال، الأسئلة الموضوعية (كالاختيار المتعدد) تصحح آليًا بدرجة واحدة لكل إجابة صحيحة، الأسئلة المفتوحة يتم تقييمها وفق Rubric مختصر يحدد عناصر الإجابة النموذجية؛ على سبيل المثال، إذا كان السؤال يتطلب ذكر فرضية صحيحة لمشكلة معطاة، تُمنح الدرجة كاملة إذا صاغ التلميذ فرضية واضحة وقابلة للاختبار ومنطقية، ونصف الدرجة إذا كانت الفرضية ناقصة أو غير واضحة تمامًا، وصفر إذا كانت الإجابة بعيدة تمامًا عن الفرضية المتوقعة، مجموع الدرجة الكلية للاختبار الورقي سيكون من 40 (بواقع درجة لكل فقرة).

الترقيم الدولي الموحد للطباعة ISSN: 3009-612X
الترقيم الدولي الموحد الالكتروني E. ISSN: 3009-6146
المجلة معرفة على قاعدة المجلات العلمية فى بنك المعرفة المصرى
وقاعدة بيانات Edu Search دار المنظومة

2 . المهام الأدائية العملية (يتم تصحيحها من خلال مقياس تقدير متدرج)

المهام الأدائية هي أنشطة أو تجارب عملية قصيرة يؤديها التلميذ تحت ملاحظة المقيم، بهدف تقييم مهارات التفكير العلمي في سياق تطبيقي، تم تصميم عدة مهام تتناسب مع مرحلة الصف السادس، وتغطي كل مهارة من المهارات المستهدفة.

جميع المهام الأدائية مصممة بحيث تجرى بشكل فردي وفي زمن قصير (10-15) دقيقة لكل مهمة، ومزودة بإرشادات واضحة للتلميذ تضمن فهم المطلوب دون إفصاح مباشر عن المهارة قيد التقييم، كما تم إعداد قائمة تقدير متدرجة خاصة بكل مهمة، يستخدمها المقيم لتسجيل أداء التلميذ على الخطوات أو المعايير المهمة، هذه المهام تضفي طابعًا عمليًا على التقييم، مما يجعل التلميذ في موقف مشابه للاستقصاء العلمي الحقيقي، ويعطي فرصة لملاحظة سلوكيات لا تظهر في الاختبارات الكتابية.

لكل مهمة أدائية تم إعداد مقياس تقدير متدرج Rubric تفصيلي يتضمن 5 معايير الأداء و3 مستويات الإتقان (إتقان واضح – إتقان جزئي – لم يُظهر المهارة)، وتتراوح الدرجة فيها بين (2 – 1 – 0)، وهكذا تحصل كل مهمة على (10) درجات، ويكون إجمالي الدرجة الكلية على المهام الأدائية يتراوح بين (0 – 80).

3 . بطاقة الملاحظة:

بطاقة الملاحظة هي أداة تقييمية يستخدمها المقيم (المعلم أو الباحث) لتسجيل ملاحظات نوعية وكمية حول أداء التلميذ أثناء قيامه بالتجارب العملية، صُممت بطاقة الملاحظة كجدول يتضمن بنودًا محددة تمثل مؤشرات كل مهارة من مهارات التفكير العلمي، بحيث يتم التأشير على درجة إتقان التلميذ لكل مؤشر خلال الملاحظة المباشرة، تم إعداد البطاقة بعد تحليل دقيق لمكونات كل مهارة وصياغة سلوكيات قابلة للرصد تعكس تلك المهارة، فيما يلي وصف لعناصر بطاقة الملاحظة:

تم تقسيم البطاقة إلى ثمانية أجزاء توافق المهارات الثماني المستهدفة، كل جزء يحتوي على عدة مؤشرات سلوكية تعكس مظاهر المهارة.

أمام كل مؤشر هناك مقياس تقدير وصفي (متقن جدًا، متقن، بحاجة إلى تحسين) تُعطى الدرجات (3،2،1)، ولتقدير مستوى أداء التلميذ في هذا المؤشر أثناء المشاهدة، يقوم المقيم بالتأشير أو إعطاء الدرجة المناسبة بناءً على سلوك التلميذ، على سبيل المثال، إذا لاحظ المقيم أن التلميذ قدم تفسيرًا علميًا صحيحًا عند سؤاله عن نتيجة التجربة، فإنه يمنحه علامة عالية في بند "يُفسّر النتائج التي توصل إليها ربطًا بالفرضية"، أما إذا عجز عن تفسير ما حصل أو أعطى تفسيرًا عشوائيًا، فيسجل درجة منخفضة في هذا البند.

الترقيم الدولي الموحد للطباعة ISSN: 3009-612X
الترقيم الدولي الموحد الالكتروني E. ISSN: 3009-6146
المجلة معرفة على قاعدة المجلات العلمية فى بنك المعرفة المصرى
وقاعدة بيانات Edu Search دار المنظومة

نتائج الدراسة ومناقشتها:

النتائج الخاصة بالفرض الأول:

ينص الفرض الأول على أنه " تتحقق لبطارية قياس مهارات التفكير العلمي (الاختبار التحصيلي – المهام الأدائية – بطاقة الملاحظة) مؤشراتٌ سيكومترية مقبولة ودالّة إحصائيًا "

وقد قامت الباحثة بالتحقق من الخصائص السيكومترية لبطارية قياس مهارات التفكير العلمي كما يلي:

1- الاختبار التحصيلي:

أ- تحليل فقرات الاختبار التحصيلي لحساب معاملات السهولة والصعوبة والتمييزية:

تم تحليل درجات عينة التحقق من الخصائص السيكومترية على الاختبار التحصيلي، وذلك بهدف حساب معاملات السهولة والصعوبة والتمييز لفقرات الاختبار وعددها (40) فقرة، وأوضحت النتائج أن معاملات الصعوبة لفقرات الاختبار تراوحت بين: (0.24- 0.76)، وهي قيم تقع فى المستوى المقبول من الصعوبة، وتتراوح قيم معاملات التمييز لفقرات الاختبار تراوحت بين: (0.2- 0.8)، وهي قيم تقع فى المستوى المقبول من التمييز، وعلى ذلك فقد تم الإبقاء على جميع مفردات الاختبار التحصيلي.

ب- صدق وثبات الاختبار التحصيلي:

تم التحقق من صدق الاختبار التحصيلي بطريقة صدق المحتوى حيث تم عرض الصورة الأولية من الاختبار التحصيلي على عدد من المحكمين من ذوي الخبرة والاختصاص بلغ عددهم (10) محكمًا، وقد اعتمدت الباحثة على حساب النسب المئوية لاتفاق المحكمين ومعامل لاوشي لصدق المحتوى فى قبول مفردات الاختبار، وكانت نسب اتفاق المحكمين على مفردات الاختبار، تتراوح بين (80 – 90) %، وتراوحت قيم معامل لاوشي لصدق المحتوى بين (0.6 – 0.8) وهي قيم تدل على صدق المحتوى للاختبار التحصيلي.

ت- ثبات الاختبار:

الثبات بطريقة ألفا كرونباخ: (Alpha Cronbach's)

تم استخدام معامل ألفا كرونباخ لحساب ثبات أبعاد الاختبار وكذلك ثبات الاختبار ككل، وقد بلغت قيمة معاملات الثبات لأبعاد الاختبار بطريقة "ألفا كرونباخ" على الترتيب: (0.720)، (0.735)، (0.715)، (0.773)، (0.751)، (0.745)، (0.750)، (0.752) كما بلغ معامل الثبات العام للاختبار (0.908)، وتؤكد جميع هذه القيم على أن الاختبار التحصيلي للجوانب المعرفية لمهارات التفكير العلمي يتمتع بدرجة عالية من الثبات.

الترقيم الدولي الموحد للطباعة ISSN: 3009-612X
الترقيم الدولي الموحد الالكتروني E. ISSN: 3009-6146
المجلة معرفة على قاعدة المجلات العلمية فى بنك المعرفة المصرى
وقاعدة بيانات Edu Search دار المنظومة

2- المهام الأدائية:

صدق المحتوى للمهام الأداء:

قامت الباحثة بعرض المهام على مجموعة من الخبراء من اساتذة مناهج وطرق تدريس العلوم وموجهين مادة العلوم بالتربية والتعليم، وعددهم (10) للتحقق من صدق المهام الأدائية في تقدير التفكير العلمي في مادة العلوم للصف السادس الابتدائي ، من حيث سلامة ووضوح الصياغة اللغوية للمواقف، ومدى مناسبة المواقف لعينة الدراسة ومرحلتها العمرية، وكذلك مدى ملائمة المهام الأدائية المحددة لقياس كل مهارة من مهارات التفكير العلمي لدى تلاميذ المرحلة الابتدائية ، وقد تراوحت نسب الاتفاق على المهام الأدائية من (70 : 90) % ، وهي تعد نسب مقبولة تدل على صلاحية المهام الأدائية ، وقامت الباحثة بعمل التعديلات المطلوبة في ضوء أراء السادة المحكمين.

3- مقياس التقدير المتدرج:

أ- صدق المحتوى:

قامت الباحثة بعرض مقياس التقدير المتدرج الذي يتم تصحيح المهام الأدائية في ضوءه على مجموعة من المحكمين وعددهم (10) المتخصصين في المناهج وطرق تدريس العلوم وعلم النفس التربوي للتحقق من تمثيل المهارات الجزئية للمهارة الأساسية المستهدف من قياسها من خلال كل مهمة أدائية، وقد تراوحت نسب الاتفاق على المهارات الجزئية لكل مهارة فرعية من (80 :90) %

ب- ثبات بطاقة الملاحظة:

تم حساب ثبات بطاقة الملاحظة من خلال أسلوب تعدد الملاحظين على أداء التلميذ الواحد، ثم حساب معامل الاتفاق بين تقديرهم للأداء، وتمت الاستعانة بأحد الزميلات، وعرض بطاقة الملاحظة عليها وتعليمات استخدامها، ثم تطبيق البطاقة على أربعة تلاميذ وتم حساب نسبة الاتفاق لكل طالب باستخدام معامل اتفاق كوبر، كما يتضح من جدول (3).

جدول **(3)**: نسبة الاتفاق بين الملاحظين على أداء الطلاب الأربعة على مقياس التقدير المتدرج لقياس أداء الطلاب على المهام الأدائية

نسبة الاتفاق للطالب الأول	نسبة الاتفاق للطالب الثاني	نسبة الاتفاق للطالب الثالث	نسبة الاتفاق للطالب الرابع	المتوسط
%86	%91	%85	%88	%80

ويتضح من جدول (3) أن بطاقة الملاحظة على درجة عالية من الثبات.

4- بطاقة ملاحظة مهارات التفكير العلمي:

أ- صدق المحتوى:

قامت الباحثة بعرض بطاقة الملاحظة على مجموعة من المحكمين وعددهم (10) المتخصصين في المناهج وطرق تدريس العلوم وعلم النفس التربوي للتحقق من تمثيل المهارات الجزئية للمهارة الأساسية المستهدف من قياسها من خلال كل

الترقيم الدولي الموحد للطباعة ISSN: 3009-612X
الترقيم الدولي الموحد الالكتروني E. ISSN: 3009-6146
المجلة معرفة على قاعدة المجلات العلمية فى بنك المعرفة المصرى
وقاعدة بيانات Edu Search دار المنظومة

مهمة أدائية، وقد تراوحت نسب الاتفاق على المهارات الجزئية لكل مهارة فرعية من (70 :90) %.

أ- ثبات بطاقة الملاحظة:

اشترك (5) من المعلمين الخبراء في تدريس العلوم للمرحلة الابتدائية فى ملاحظة أداء التلاميذ أثناء التجارب العملية لدى عدد (10) تلاميذ ، وسجل كل منه مُلاحظاته فى ضوء عبارات بطاقة الملاحظة ، وتم استخدام معادلة كوبر لحساب نسبة الاتفاق بين الملاحظين حيث بلغت قيمة معاملات الثبات لمحاور بطاقة الملاحظة على الترتيب: (91.6%)؛ (93.6%)؛ (96%)؛ (97.5%) ؛ (97.5%) ؛ (94.2%) ؛ (95%) ؛ (96.6%) وهى قيم تؤكد على أن محاور بطاقة تقييم المنتج بدرجة مرتفعة من الثبات ، وبلغت قيمة معامل الثبات العام للبطاقة ككل (95.25%)، وهى قيمة تؤكد على أن بطاقة الملاحظة تتمتع بدرجة مرتفعة من الثبات.

ويتضح مما سبق أن الفرض الأول قد تحقق حيث كانت قيم مؤشرات الصدق والثبات لكل من (الاختبار التحصيلي – المهام الأدائية – بطاقة الملاحظة) مقبولة ودالّة إحصائيًا.

النتائج الخاصة بالفرض الثاني:

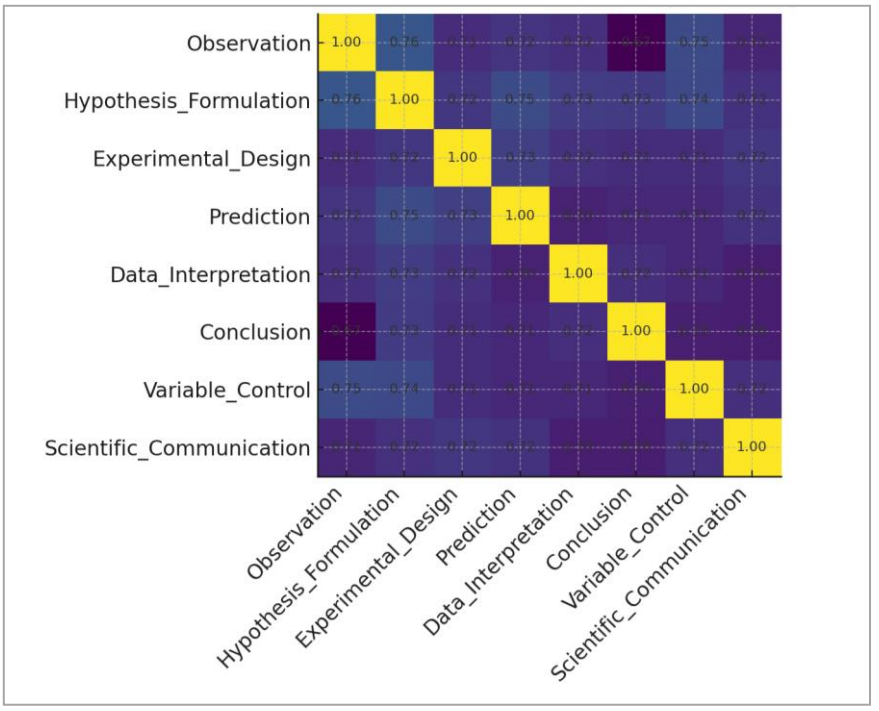
ينص الفرض الثاني على أنه " يتمتع نموذج تعلم الاله **المستخدم (Random Forest)** بقدرة تصنيفية دالة إحصائيًا في تحديد مستوى التفكير العلمي لدى التلاميذ (ضعيف – متوسط-متفوق) بنسبة دقة لا تقل عن **80%** "

أولاً: الإحصاءات الوصفية:

تم توحيد درجات المهارات الفرعية للتفكير العلمي ـــ الملاحظة، صياغة الفرضيات، التصميم التجريبي، التنبؤ، تفسير البيانات، الاستنتاج، ضبط المتغيرات، والتواصل العلمي ـــ على مقياسٍ موحّد واحد؛ بما يسهّل المقارنة بين الأبعاد المختلفة وتجميعها في درجة مركّبة تمثّل الأداء الكلي للتفكير العلمي لدى التلاميذ ، على مقياس موحّد يتيح المقارنة بين الأبعاد المختلفة، ثم حساب المؤشرات الوصفية الأساسية (المتوسط، الوسيط، الانحراف المعياري، المدى الربيعي) للتحقق من اتساق التوزيعات ورصد أي قيم متطرفة محتملة، ويأتي بعد ذلك فحص العلاقات البينية بين المهارات عبر مصفوفة الارتباط بوصفها خطوة لازمة لأمرين:

- التحقق البنائي الأولي من كون المهارات تعمل كمنظومة مترابطة تمثّل بُنية "التفكير العلمي".
- تكوين درجة مركبة تعكس الأداء الكلي، مع الاطمئنان في الوقت نفسه لعدم وجود تعارضات قوية (Multicollinearity) قد تُضعف تقدير النماذج التنبؤية لاحقًا.

الترقيم الدولي الموحد للطباعة ISSN: 3009-612X
الترقيم الدولي الموحد الالكتروني E. ISSN: 3009-6146
المجلة معرفة على قاعدة المجلات العلمية فى بنك المعرفة المصرى
وقاعدة بيانات Edu Search دار المنظومة



يعتمد هذا الفحص على معاملات ارتباط خطي (معامل بيرسون) بين كل زوج من المهارات، ويُعرض بصريًا في صورة خريطة حرارية تُظهر قوة الاتجاه الموجب/السالب وشدّته، بما يسهّل قراءة النمط العام للعلاقات وتحديد الأبعاد الأكثر ارتباطاً، ويوضح شكل (1) ارتباطات موجبة مرتفعة بين معظم المهارات الفرعية، مما يبرر تكوين "درجة مركبة" واستخدام نماذج متعددة المتغيرات.

شكل (1) مصفوفة الارتباط بين مهارات التفكير العلمي

تُظهر المصفوفة تراكبًا ملحوظًا بين المهارات—وخاصة الملاحظة (Observation)، وصياغة الفرضيات (Hypothesis Formulation)، وتفسير البيانات (Data Interpretation) —بما ينسجم مع كونها أبعادًا متساندة في بناء التفكير العلمي. هذا النمط يدعم صلاحية التجميع في مؤشر مركّب (Composite Index)، ويشير إلى أن تمييز الفئات (ضعيف – متوسط – ممتاز) ينبغي أن ينعكس عبر مجموع المهارات لا مهارة منفردة، كما يطمئن إلى غياب تعارضات شديدة (Multicollinearity) قد تُضعف استقرار النمذجة اللاحقة (Model Stability).

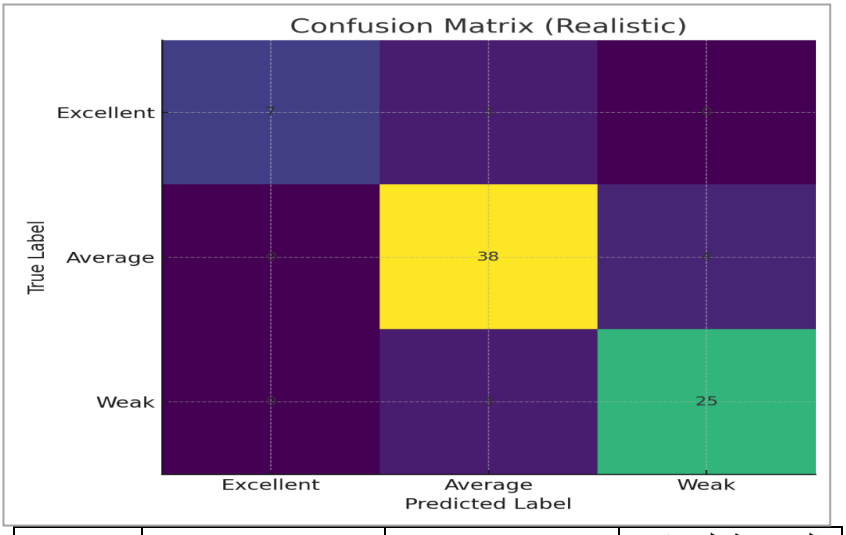
ثم قامت الباحثة بتقييم أداء نموذج الغابة العشوائية باستخدام مصفوفة الالتباس ومقاييس الدقة والاستدعاء ومعامل F1 وأظهرت النتائج أن النموذج حقق دقة

الترقيم الدولي الموحد للطباعة ISSN: 3009-612X
الترقيم الدولي الموحد الالكتروني E. ISSN: 3009-6146
المجلة معرفة على قاعدة المجلات العلمية فى بنك المعرفة المصرى
وقاعدة بيانات Edu Search دار المنظومة

تصنيف كلية 87.5%، أي أعلى من 80%، كما يتضح من جدول (4)، ويُظهر شكل (2) مصفوفة الالتباس لتوزيع الحالات المتوقع والناتج حيث (المحور العمودي: الفئة الحقيقية، الأفقي: الفئة المتنبَّأ بها)

جدول (4) تقرير تصنيف الغابة العشوائية

الفئة	الدقة (Precision)	الاستدعاء (Recall)	درجة F1
متوسط	0.864	0.905	0.884
ممتاز	1.000	0.700	0.824
ضعيف	0.862	0.893	0.877
الدقة الكلية (Accuracy)	—	—	0.875
المتوسط (ماكرو)	0.909	0.833	0.861
المتوسط المرجّح	0.880	0.875	0.874



شكل (2) مصفوفة الالتباس لنموذج الغابة العشوائية.

الترقيم الدولي الموحد للطباعة ISSN: 3009-612X
الترقيم الدولي الموحد الالكتروني E. ISSN: 3009-6146
المجلة معرفة على قاعدة المجلات العلمية فى بنك المعرفة المصرى
وقاعدة بيانات Edu Search دار المنظومة

تظهر النتائج دقة عالية في تمييز مستويات التفكير العلمي الثلاثة، حيث بلغت الدقة الإجمالية 0.875، مما يؤكد تحقق الفرضية الثانية بتحقيق دقة تصنيف أعلى من 80%

توضح مصفوفة الالتباس تركزًا قويًا على القطر الرئيس، ما يعني أن معظم الحالات صُنّفت في فئاتها الصحيحة (Excellent/ Average/ Weak) الانحرافات خارج القطر محدودة وتعكس التداخل الطبيعي بين "متوسط" وباقي الفئات، هذا النمط يدعم تحقق الفرض الثاني بأن النموذج يمتلك قدرة تصنيفية دالة إحصائيًا ومجديًا عمليًا.

تُعزى هذه الدقة العالية إلى قدرة الغابة العشوائية على دمج قرارات أشجار متعددة لتقليل الأخطاء .(majority voting) وقد أشارت دراسات مماثلة إلى أن نماذج الغابة العشوائية غالباً ما تتفوق على خوارزميات فردية في مهام التصنيف بسبب هذا التجميع.(Donges, 2024)

النتائج الخاصة بالفرض الثالث:

ينص الفرض الثالث على أنه " توجد علاقة ارتباطيه موجبة ودالة إحصائيًا بين متوسط درجات التلاميذ في مهارات التفكير العلمي والتصنيف الناتج عن نموذج تعلم الاله "

يستهدف هذا الفرض التحقّق من اتجاه وقوة العلاقة بين الأداء الكلي في مهارات التفكير العلمي (الدرجة المركّبة) والمستوى التصنيفي النهائي للتلميذ (ضعيف – متوسط – ممتاز)، نظرًا لطبيعة الفرض (علاقة خطية متوقعة واتجاه موجب)، فقد تم توظيف معامل الارتباط لبيرسون لقياس شدة واتجاه العلاقة بين متغيرين مستمرَّين، مع اعتماد ترميزٍ عدديٍّ منظم لمتغير التصنيف (ضعيف=1، متوسط=2، ممتاز=3) بما يعكس ترتيب المستويات تصاعديًا.

ويُستخدم معامل ارتباط بيرسون عندما يكون الهدف تقدير قوة علاقة خطية بين متغير كمي (الدرجة المركبة) ومتغير ترتيبي مُرَقّم على نحو يعكس رُتَب التصنيف، مع توافر حجم عينة كافٍ وغياب القيم الشاذة المؤثرة بدرجة مفرطة.

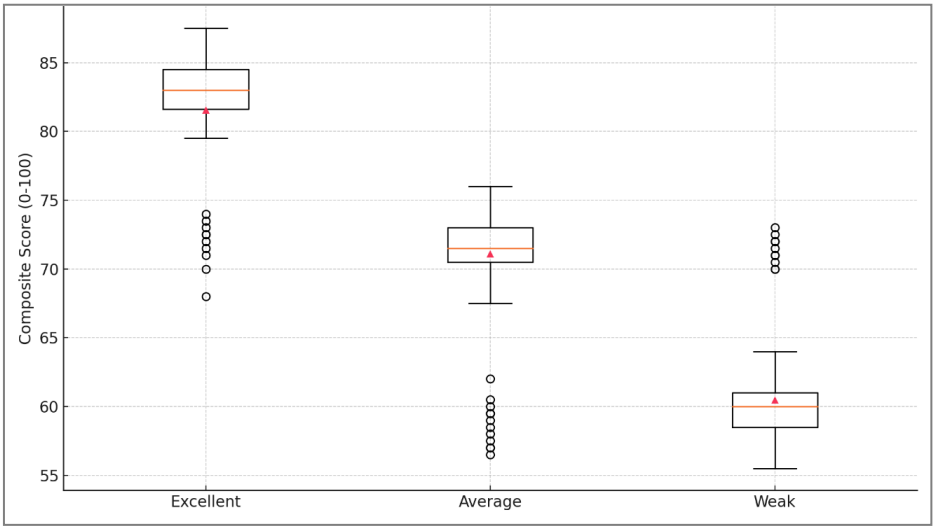
إجراءات التحليل والترميز:

- الدرجة المركبة: ناتج تجميع الدرجات الموحَّدة عبر الأبعاد الفرعية للتفكير العلمي بعد توحيد المقاييس.
- ترميز التصنيف: ضعيف=1، متوسط=2، ممتاز=3 (ترميز أحادي البُعد يعكس الترتيب المنطقي للمستويات)، ومن ثم تم حساب معامل ارتباط بيرسون (r)وقيمة p للدلالة الإحصائية.

الترقيم الدولي الموحد للطباعة ISSN: 3009-612X
الترقيم الدولي الموحد الالكتروني E. ISSN: 3009-6146
المجلة معرفة على قاعدة المجلات العلمية فى بنك المعرفة المصرى
وقاعدة بيانات Edu Search دار المنظومة

وأوضحت النتائج وجود ارتباط إيجابي قوي يبلغ حوالي$r = 0.88$ ، مع دلالة إحصائية عالية .($p \approx 0$) ، ويُعني هذا أن التلاميذ ذوو التصنيف الأعلى يميلون إلى الحصول على درجات تراكمية أكبر في مهارات التفكير العلمي، هذا يتوافق مع مفهوم الارتباط الإيجابي في التحليل الإحصائي حيث يزداد أداء الطلاب بزيادة المستوى التصنيفي (Cohen et al., 2003)، وتدعم الأدبيات التربوية مثل دراسة فتحية اللولو (1997) وجود علاقة ارتباطية دالة بين المهارات العلمية وتحصيل الطلاب (أي أن زيادة مستوى المهارة ترافقها زيادة في الأداء العام)، مما يتفق مع النتائج الحالية.

ويُعد الشكل (3) خطيًا بصرياً للعلاقة عبر مخططات الصندوق Box Plot، والشكل (4) مخطط الإسقاط بالمكوّنات الرئيسة (PCA) كأدلة مسانِدة على اتجاه الارتباط وتميّز المجموعات.

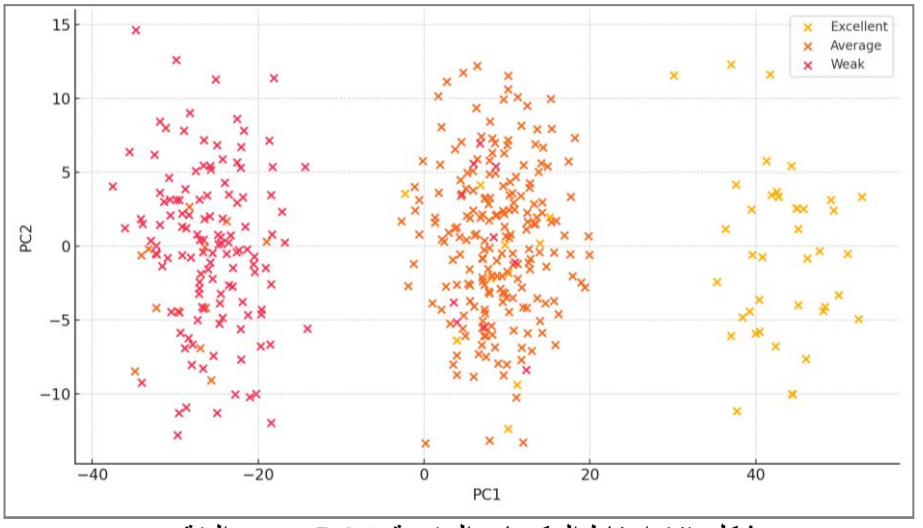


شكل (3) Boxplot للدرجة المركبة حسب الفئة (Composite Score (by Class

يعرض شكل (3) تدرجًا صاعدًا واضحًا في الوسيط والارباعيات من Weakإلى Excellent ، بما يدعم وجود علاقة موجبة قوية بين الدرجة المركبة والتصنيف، ويدل ارتفاع الوسيط وازدياد المجال بين الارباعيات مع الانتقال من Weak → Average → Excellentيؤكد أن التلاميذ ذوي التصنيف الأعلى يحققون درجات مركبة أعلى بصورة منهجية، وهو تمثيل بصري مباشر للعلاقة الارتباطية الموجبة المفحوصة إحصائيًا في هذا الفرض، هذا النمط يعزّز

الترقيم الدولي الموحد للطباعة ISSN: 3009-612X
الترقيم الدولي الموحد الالكتروني E. ISSN: 3009-6146
المجلة معرفة على قاعدة المجلات العلمية فى بنك المعرفة المصرى
وقاعدة بيانات Edu Search دار المنظومة

الاستدلال الإحصائي العددي (r = 0.88) ويشير إلى انتظام العلاقة عبر المجموعات الثلاث، مع تداخل محدود متوقع في مجموعة متوسط.



شكل (4) إسقاط المكوّنات الرئيسة PCA بحسب الفئة

يؤكد شكل (4) على وجود انفصالٍ معقول بين المجموعات على المستويين PC1وPC2، بما يتسق مع قوة العلاقة بين نمط الدرجات والتصنيف، حيث يُظهر الإسقاط تمركز نقاط فئة Excellent في حيّز مغاير عنWeak ، مع تراكب محدود معAverage ، مما يترجم بصريًا إلى وجود بُعد (أو أكثر) يلتقط الفروق المرتبطة بارتفاع الدرجة، وتؤكد كل من البيانات الرقمية والدلائل البصرية على وجود علاقة ارتباطية موجبة قوية بين الدرجة المركبة والتصنيف، بما يدعم قبول الفرض الثالث.

النتائج الخاصة بالفرض الرابع:

ينص الفرض الرابع على أنه " توجد فروق ذات دلالة إحصائية بين متوسط درجات التلاميذ في المهارات العلمية تبعاً لمستوى تصنيفهم النهائي (ضعيف – متوسط – ممتاز)."

تم تطبيق تحليل التباين أحادي الاتجاه (One-way ANOVA) على درجات المهارات الموحّدة عبر ثلاث مجموعات: ضعيف، متوسط، ممتاز، كما تم حساب حجم الأثر (η^2) لتقدير الدلالة العملية، ومعرفة اتجاه الفروق بين التباينات (Cohen, 1988)، كما أُجري اختبار Tukey HSD للمقارنات البعدية، للتعرّف على أزواج المجموعات التي تختلف دلاليًا عند كل مهارة (Lawson, 2000).

الترقيم الدولي الموحد للطباعة ISSN: 3009-612X
الترقيم الدولي الموحد الالكتروني E. ISSN: 3009-6146
المجلة معرفة على قاعدة المجلات العلمية فى بنك المعرفة المصرى
وقاعدة بيانات Edu Search دار المنظومة

جدول (5) تحليل التباين أحادي الاتجاه حسب المستوى النهائي (لكل مهارة)

المهارة	قيمة F	P-Value	حجم الأثر η^2
الملاحظة	312.80	< .001	0.612
الاستنتاج	300.92	< .001	0.603
التنبؤ	282.02	< .001	0.587
تصميم التجربة	281.76	< .001	0.587
تفسير البيانات	278.09	< .001	0.584
التواصل العلمي	269.78	< .001	0.576
صياغة الفرضية	259.75	< .001	0.567
ضبط المتغيرات	253.00	< .001	0.560

من النتائج الموضحة بجدول (5) يتضح أن جميع القيم **دالّة بقوة**، وقيم حجم الأثر كبيرة إلى متوسطة ($\eta^2 \approx 0.56–0.61$)، ما يعني أن عامل "المستوى النهائي" يفسّر نسبة كبيرة من تباين الدرجات في كل المهارات —أي أن التلاميذ المتميزين يتخطّون المتوسطين، والمتوسطين يتخطّون الضعفاء بشكل منسّق عبر الأبعاد (Cohen, 1988; Lawson, 2000). ويوضح جدول (6) المتوسطات الحسابية والانحرافات المعيارية للدرجة المركّبة حسب مستوى مهارات التفكير العلمي.

جدول (6) الإحصاءات الوصفية للدرجة المركّبة حسب المستوى

مستوى التفكير العلمي	العدد(n)	المتوسط	الانحراف المعياري
ممتاز	50	81.54	5.23
متوسط	212	71.08	3.44
ضعيف	138	60.47	3.46

ويتضح من جدول (6):

- الفروق بين المتوسطات كبيرة جدًا ومنتظمة :ممتاز > (≈81.5) متوسط > (≈71.1)ضعيف (≈60.5)، الفاصل بين كل مستويين يقارب 11–10 نقطة، وهو فارق تعليمي واضح على مقياس الدرجة المركّبة.
- التشتت داخل كل فئة منخفض نسبيًا (SD≈3.4–5.2)، ما يعني أن الأداء متجانس داخل المجموعة الواحدة؛ وهذا يعزّز ثبات الفروق بين المجموعات.
- حجم الأثر الزوجي (Cohen's d):
  - ممتاز مقابل متوسط $d \approx 2.72$ أثر كبير
  - متوسط مقابل ضعيف $d \approx 3.08$ أثر كبير
  - ممتاز مقابل ضعيف $d \approx 5.27$ أثر كبير جدًا

الترقيم الدولي الموحد للطباعة ISSN: 3009-612X
الترقيم الدولي الموحد الالكتروني E. ISSN: 3009-6146
المجلة معرفة على قاعدة المجلات العلمية فى بنك المعرفة المصرى
وقاعدة بيانات Edu Search دار المنظومة

هذه القيم تؤكّد أن الفروق ليست إحصائية فحسب، بل ذات معنى تربوي قوي (Cohen, 1988).

كما تم إجراء الاختبارات البعدية (Tukey HSD) على الدرجة المركّبة، حيث أوضحت النتائج أن جميع المقارنات الثنائية دالّة إحصائيًا عند مستوى (0.01) ممتاز مقابل متوسط، متوسط مقابل ضعيف، ممتاز مقابل ضعيف، ما يثبت التمايز المنهجي بين المستويات.(Lawson, 2000)

وبالتالي فإنه هناك فاصل حقيقي وكبير بين المستويات الثلاثة في الأداء العام، مدعوم بإحصاءات وصفية واضحة وأحجام أثر هائلة واختبارات بعدية توضح وجود فروق دالة إحصائيًا هذه الصورة تؤكّد صلاحية التصنيف النهائي وتمييزه الواقعي بين المتعلمين.(Lawson, 2000)

كما تم فحص "بروفايل" كل مهارة على حدة عبر المستويات بغرض الإجابة عن السؤال "هل الفروق عامة عبر البنية المهارية كلها، أم مرتبطة ببعض الأبعاد فقط؟"، ويوضح جدول (7) المتوسطات الحسابية للمهارات الموحدة حسب مستوى مهارات التفكير العلمي، لكل مهارة من المهارات الفرعية.

جدول (7) متوسط المهارات المُوحَّدة حسب المستوى(N=400)

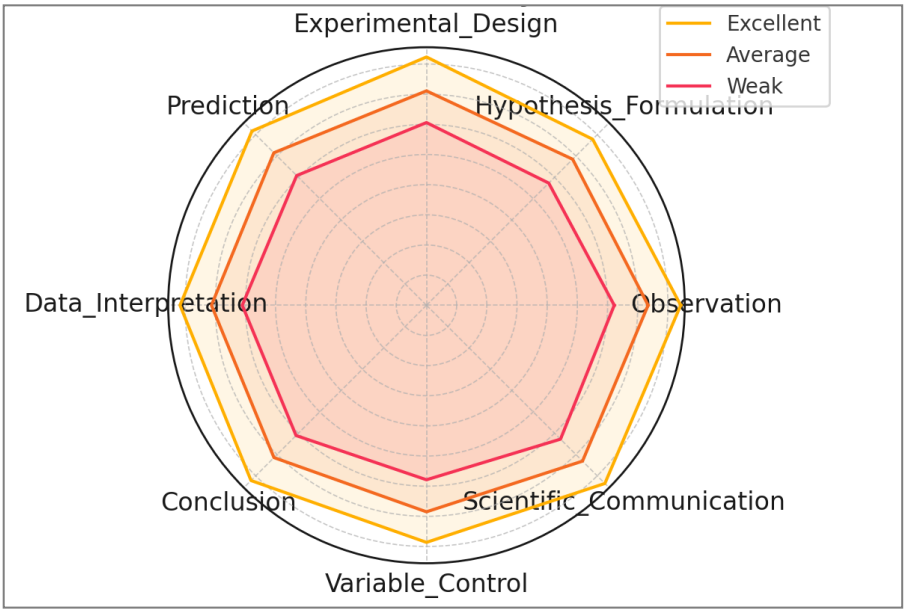
المهارة	ضعيف	متوسط	ممتاز
الملاحظة	62.21	73.43	84.22
صياغة الفرضية	57.30	68.49	77.87
تصميم التجربة	60.59	71.10	82.34
التنبؤ	60.82	71.51	81.74
تفسير البيانات	61.03	71.24	81.65
الاستنتاج	61.11	71.45	82.21
ضبط المتغيرات	57.88	68.51	78.64
التواصل العلمي	62.80	73.11	83.59

وتم إجراء مقارنات بعدية (Tukey HSD) عبر كل مهارة لجميع المهارات الثماني، كل الأزواج الثلاثية (ممتاز–متوسط، متوسط–ضعيف، ممتاز–ضعيف) حيث كانت الفروق دالة إحصائيًا عند مستوى (0.01)، حيث أوضحت النتائج وجود نمط متزايد منتظم من ضعيف ← متوسط ← ممتاز عبر جميع المهارات يدعم فرضية البنية المشتركة للتفكير العلمي في هذه المرحلة (Kuhn, 2010; Zimmerman, 2007،

الترقيم الدولي الموحد للطباعة ISSN: 3009-612X
الترقيم الدولي الموحد الالكتروني E. ISSN: 3009-6146
المجلة معرفة على قاعدة المجلات العلمية فى بنك المعرفة المصرى
وقاعدة بيانات Edu Search دار المنظومة

ويتضح من جدول (7):

- النمط **متزايد ومتّسق** في جميع الأبعاد: كل مهارة ترتفع من ضعيف إلى متوسط إلى ممتاز دون استثناء—وهذا **انتظام بنيوي** قوي.
- **أكبر فجوتين** بين ممتاز وضعيف ظهرتا في:
  - **الملاحظة** :فرق **22.01** ≈ نقطة(62.21 – 84.22)
  - **تصميم التجربة** :فرق **21.75** ≈ نقطة (60.59 – 82.34) يليهما الاستنتاج، التنبؤ، تفسير البيانات، التواصل العلمي بفروق **20–21** ≈نقطة، ثم صياغة الفرضية وضبط المتغيرات بفروق **20–21** ≈نقطة أيضًا.



شكل (5) رادار متوسطات المهارات الموحدة حسب الفئة **(Radar – Mean Skills)**

ويوضح شكل (5) اتساع "بصمة" الأداء من Weak إلى Excellent عبر جميع المهارات، متسقًا مع دلالة فروق ANOVA.

يُظهر الرادار اتساق التفوق في كل محور مهاري لفئة Excellent يليها Average ثمWeak ، ما يوفر تصورًا بديهيًا لفروق المتوسطات التي أكدها تحليل التباين، هذا التوافق بين الدليلين الجدولي (ANOVA) والبصري (الرادار) يعزّز قوة الاستنتاج بوجود فروق حقيقية بين المجموعات.

الترقيم الدولي الموحد للطباعة ISSN: 3009-612X
الترقيم الدولي الموحد الالكتروني E. ISSN: 3009-6146
المجلة معرفة على قاعدة المجلات العلمية فى بنك المعرفة المصرى
وقاعدة بيانات Edu Search دار المنظومة

وبالتالي فإن التلاميذ الأقوى لا يبرعون فقط في الاستدلال النهائي، بل يبدأ تفوقهم مبكرًا من جودة الملاحظة وأدوات التصميم التجريبي لبنات أساسية تدعم كل السلسلة اللاحقة.(Zimmerman, 2007; Lawson, 2000)

خلاصة بروفايل المهارات :التفوق ينتظم عبر الملاحظة ← الفرضية ← التصميم ← التنبؤ ← تفسير البيانات ← الاستنتاج ← ضبط المتغيرات ← التواصل؛ نمط "متوازٍ" يدعم فرضية البنية المشتركة لمهارات التفكير العلمي في نهاية الابتدائي.(Kuhn, 2010; Zimmerman, 2007)

النتائج الخاصة بالفرض الخامس:

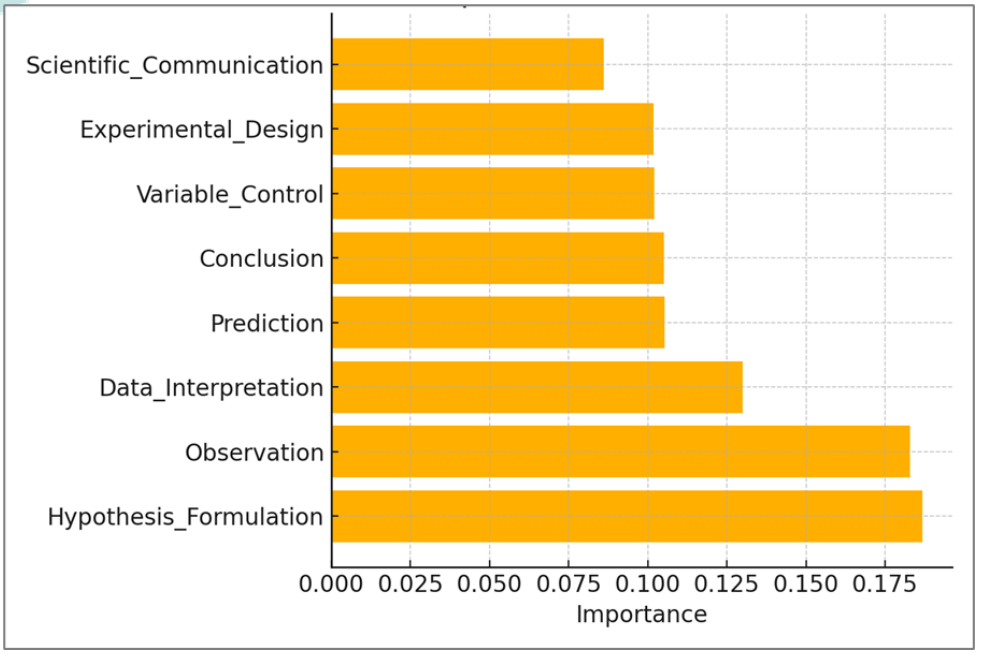
ينص الفرض الخامس على أنه "تسهم بعض المهارات الفرعية بدرجة أكبر من غيرها في التنبؤ بمستوى التفكير العلمي لدى التلاميذ"

لمعرفة أي المهارات تضيف أكثر إلى دقة التصنيف، اعتمدنا على أهمية الميزات المحسوبة من نموذج الغابة العشوائية. يمكن استخدام مقياس Gini أو خفض قيمة النقاء (impurity) لهذا الغرض، حيث أظهرت النتائج أن مهارة صياغة الفرضية كانت الأكثر تأثيراً (أهمية ≈0.187)، تليها مهارة الملاحظة العملية (0.183)، ثم مهارات التفسير البياني والتنبؤ والتوصل والاستنتاج والقضايا التجريبية بترتيب تنازلي حتى الاتصال العلمي الأقل (0.086) كما فيما يلي ، وهذا يعني أن بعض المهارات الفرعية تساهم بنسبة أكبر في تحديد مستوى التفكير العلمي.

- صياغة الفرضية: الأهمية 18.68%
- الملاحظة العملية: الأهمية 18.28%
- التفسير البياني: الأهمية 13.00%
- التنبؤ: الأهمية 10.54%
- الاستنتاج: الأهمية 10.52%
- ضبط المتغيرات: الأهمية 10.20%
- التصميم التجريبي: الأهمية 10.18%
- الاتصال العلمي: الأهمية 8.61%

استخدام الغابة العشوائية يسهل استنتاج مثل هذه النتائج، فقد صُممت الخوارزمية أساساً لتحديد أهمية كل متغير تنبؤي وتؤكد هذه النتيجة فرضية البحث بأن بعض المهارات الفرعية (مثل الفرضية والملاحظة) لها وزن أكبر في نموذج التنبؤ من غيرها، وقد أشار Donges (2024) إلى أن ميزة قوة الغابة العشوائية تتمثل في سهولة استخراج مؤشرات أهمية المتغيرات لتفسير النموذج، مما يجعلها مناسبة لهذا النوع من التحليل التربوي، ويبيّن شكل (6) ترتيب المهارات من الأعلى تأثيرًا إلى الأقل في التنبؤ بمستوى التفكير العلمي.

الترقيم الدولي الموحد للطباعة ISSN: 3009-612X
الترقيم الدولي الموحد الالكتروني E. ISSN: 3009-6146
المجلة معرفة على قاعدة المجلات العلمية فى بنك المعرفة المصرى
وقاعدة بيانات Edu Search دار المنظومة



شكل (6) أهمية السمات (Feature Importance) في نموذج Random Forest

ويوضح الشكل (6) أن صياغة الفرضيات والملاحظة تأتي في الصدارة، تليهما تفسير البيانات، ثم مجموعة مهارات التنبؤ والاستنتاج وضبط المتغيرات والتصميم التجريبي، وأخيرًا التواصل العلمي .يشير هذا التدرّج إلى أن بناء الفرضيات والملاحظة العملية يمثلان "محرّكين" تنبؤيين أساسيين، ما يوجِّه التوصيات التربوية إلى تركيز التدريب عليهما دون إغفال بقية الأبعاد.

تعقيب عام على النتائج:

- **القدرة التصنيفية العالية (Accuracy=87.5%)** تؤكد قابلية توظيف التعلّم الآلي لتقييم مهارات العلم في الابتدائي حين تتوافر بيانات غنيّة ومتعددة المصادر (Romero & Ventura, 2020; Baker & Yacef, 2009).
- **الارتباطات القوية** بين متوسط المهارات والتصنيف تعكس «بنية عامة» للتفكير العلمي تتكوّن من تآزر الملاحظة–الفرضية–التجريب–التفسير–الاستنتاج، كما تقرّه أدبيات التفكير العلمي المبكّر (Zimmerman, 2007; Kuhn, 2010).

الترقيم الدولي الموحد للطباعة ISSN: 3009-612X
الترقيم الدولي الموحد الالكتروني E. ISSN: 3009-6146
المجلة معرفة على قاعدة المجلات العلمية فى بنك المعرفة المصرى
وقاعدة بيانات Edu Search دار المنظومة

- **توجد فروق كبيرة ودالّة η^2 كبيرة** بين المستويات عبر جميع المهارات تدعم صلاحية التصنيف وبُعده البنائي (Lawson, 2000; Padilla, 1990) .
- **أكثر المهارات إسهامًا في التنبؤ** كانت «الفرضية» و«الملاحظة» ثم «تفسير البيانات»، بما يتسق مع نماذج التطور المعرفي التي تجعل ضبط المتغيرات والاستدلال السببي يتبلوران تدريجيًا بنهاية المرحلة الابتدائية . (Chen & Klahr, 1999; Koerber et al., 2015)

وبشكل عام، أظهرت النتائج اتساقاً مع التوقعات والدراسات العلمية السابقة. فتفوق النموذج التصنيفي بحصوله على دقة عالية (87.5%) يتوافق مع ما ورد في دراسات تعلم الآلة التي تثبت قوة الغابة العشوائية في مهام التصنيف بفضل تجميعها لقرارات أشجار متعددة (Donges, 2024; Scikit-learn, 2023).كذلك، يؤكد الارتباط القوي بين الدرجات والتصنيف الأثر الإيجابي المباشر للمهارات على مستوى الأداء العلمي للطالب، وهو ما يشير إليه بحث فتحية اللولو (1997) وغيره من الأبحاث التي ربطت بين مستويات المهارة والتحصيل. أما الفروقات الكبيرة في مهارات كل مجموعة، فأثبتها اختبار ANOVAكمتوقع، مدعومة بالرسم الرادارى الذي يشير إلى ارتفاع أداء مجموعة المتفوقين في جميع المهارات مقارنة بالضعفاء.

وتدل أهمية الميزات المستخرجة (Feature Importance) على أن صياغة الفرضيات تعد مهارة مركزيّة في بناء التفكير العلمي للطلاب، تليها الملاحظة العملية، وهو ما قد يستدعي من المعلمين التركيز على تنمية هاتين المهارتين لتحقيق أفضل تصنيف. يشير ذلك إلى أن المهارات العليا في التفكير العلمي ليست متكافئة التأثير، وقد أجرت دراسات أخرى) مثل (Donges, 2024 دراسات تطبيقية أوثقت كيفية اختلاف مساهمة متغيرات متعددة في نموذج تنبؤي بغرض تحسين أداء النظام.

على ضوء ما سبق، يتضح أن النموذج الذكي المقترح فعال إحصائيًا في تصنيف مستويات التفكير العلمي للتلاميذ بما يتجاوز المتطلبات(H2) ، وأن ثمة ارتباطاً وثيقاً بين الأداء التصنيفي والدرجات الفعلية للتلاميذ(H3) ، كما أن المتوسطات تختلف اختلافاً جوهرياً حسب المستوى(H4) ، وأن مهارات فرعية معينة تبرز مساهمتها في التنبؤ بمستوى الطالب أكثر من غيرها .(H5) تعكس هذه النتائج إمكانية توظيف تقنيات التعلم الآلي في التقييم التربوي وتوليد رؤى جديدة مساعدة للمعلمين والمطورين التربويين، مشددةً على ضرورة اختيار الخوارزميات المناسبة مثل الغابة العشوائية التي أثبتت كفاءتها في هذا السياق.

الترقيم الدولي الموحد للطباعة ISSN: 3009-612X
الترقيم الدولي الموحد الالكتروني E. ISSN: 3009-6146
المجلة معرفة على قاعدة المجلات العلمية فى بنك المعرفة المصرى
وقاعدة بيانات Edu Search دار المنظومة

المراجع:

إلهام علي الشلبي، شذى أحمد الخليفة. (2017). مستوى مهارات التفكير العلمي والتفكير الرياضي لدى طالبات المرحلة الابتدائية. مجلة العلوم التربوية والنفسية، 1(3) , 117-142.

رياض بهجات. (2003). التعلم الاستراتيجي مدخل مقترح لحفز التفكير (H-L-W-K) في تنمية مهارات التفكير العلمي لدى طالبات الصف الثاني المتوسط في مادة الفيزياء. مجلة كلية التربية الأساسية، جامعة بابل، 13، 332-367.

عبد الحميد زيتون. (2008). أساليب تدريس العلوم (ط. 6). عمّان، الأردن: دار الشروق.

فتحية اللولو. (1997). أثر إثراء منهج العلوم بمهارات تفكير علمي على تحصيل الطلبة في الصف السابع. رسالة ماجستير غير منشورة، كلية العلوم التربوية، الجامعة الإسلامية، غزة.

Aktamis, H., & Ergin, O. (2008). The effect of scientific process skills education on students' scientific creativity, science attitudes and academic achievements. Asia-Pacific Forum on Science Learning and Teaching, 9(1), Article 4.

Baker, R. S., & Yacef, K. (2009). The state of educational data mining in 2009: A review and future visions. Journal of Educational Data Mining, 1(1), 3–17.

Bransford, J. D., Brown, A. L., & Cocking, R. R. (Eds.). (2000). How people learn: Brain, mind, experience, and school. National Academies Press.

Breiman, L. (2001). Random forests. Machine Learning, 45(1), 5–32.

Brusilovsky, P., & Millán, E. (2007). User models for adaptive hypermedia and adaptive educational systems. In P. Brusilovsky, A. Kobsa, & W. Nejdl (Eds.), The Adaptive Web: Methods and Strategies of Web Personalization (pp. 3–53). Springer.

Chen, Z., & Klahr, D. (1999). All other things being equal: Acquisition and transfer of the control of variables strategy. Child Development, 70(5), 1098–1120.

ISSN: 3009-612X الترقيم الدولي الموحد للطباعة
E. ISSN: 3009-6146 الترقيم الدولي الموحد الالكتروني
المجلة معرفة على قاعدة المجلات العلمية فى بنك المعرفة المصرى
وقاعدة بيانات Edu Search دار المنظومة

Clark, R. E., & Feldon, D. (2014). Ten common but flawed arguments for using brain research in education. In K. R. Harris, S. Graham, & T. Urdan (Eds.), APA educational psychology handbook: Vol. 3. Application to learning and teaching (pp. 157–194). American Psychological Association.

Cohen, J. (1988). Statistical power analysis for the behavioral sciences (2nd ed.). Lawrence Erlbaum.

Conati, C., & Kardan, S. (2013). Student modeling: Supporting personalized instruction, from problem solving to exploratory open-ended activities. AI Magazine, 34(3), 13–26.

Donges, N. (2024, November 26). Random Forest: A Complete Guide for Machine Learning. Built In. Retrieved from https://builtin.com/data-science/random-forest-algorithm

Duschl, R. A., Schweingruber, H. A., & Shouse, A. W. (2007). Taking Science to School: Learning and Teaching Science in Grades K-8. Washington, DC: National Academies Press.

Freeman, S., Eddy, S. L., McDonough, M., Smith, M. K., Okoroafor, N., Jordt, H., & Wenderoth, M. P. (2014). Active learning increases student performance in science, engineering, and mathematics. Proceedings of the National Academy of Sciences, 111(23), 8410–8415.

Gopnik, A. (2012). Scientific thinking in young children: Theoretical advances, empirical research, and policy implications. Science, 337(6102), 1623–1627.

Hattie, J. (2009). Visible learning: A synthesis of over 800 meta-analyses relating to achievement. Routledge.

Hopkins, K. D., Stanley, J. C., & Hopkins, B. R. (1990). Educational and Psychological Measurement and Evaluation (7th ed.). Englewood Cliffs, NJ: Prentice Hall.

ISSN: 3009-612X الترقيم الدولي الموحد للطباعة
E. ISSN: 3009-6146 الترقيم الدولي الموحد الالكتروني
المجلة معرفة على قاعدة المجلات العلمية فى بنك المعرفة المصرى
وقاعدة بيانات Edu Search دار المنظومة

Hrouzková, T., & Richterek, L. (2021). Lawson classroom test of scientific reasoning at entrance university level. In Proceedings of the 4th International Baltic Symposium on Science and Technology Education (pp. 74-85).

Hwang, G. J., & Chang, H. F. (2011). A formative assessment-based mobile learning approach to improving the learning attitudes and achievements of students. Computers & Education, 56(4), 1023–1031.

Inhelder, B., & Piaget, J. (1958). The Growth of Logical Thinking from Childhood to Adolescence: An Essay on the Construction of Formal Operational Structures. London: Routledge.

Jolliffe, I. T. (2002). Principal component analysis (2nd ed.). Springer.

Koerber, S., Osterhaus, C., & Sodian, B. (2015). Scientific thinking in elementary school: Children's conceptual knowledge and inquiry skills. Journal of Cognition and Development, 16(4), 617–641.

Koslowski, B. (1996). Theory and evidence: The development of scientific reasoning. MIT Press.

Kuhn, D. (2002). What is scientific thinking and how does it develop? In U. Goswami (Ed.), Blackwell handbook of childhood cognitive development (pp. 371–393). Blackwell.

Kuhn, D. (2010). Teaching and learning science as argument. Science Education, 94(5), 810–824.

Kuhn, D. (2010). What is scientific thinking and how does it develop? The Wiley-Blackwell handbook of childhood cognitive development, 497-523.

Kuhn, D., & Pearsall, S. (2000). Developmental origins of scientific thinking. Journal of Cognition and Development, 1(1), 113–129.

Lawson, A. E. (1978). The development and validation of the classroom test of formal reasoning. Journal of Research in Science Teaching, 15(1), 11–24.

الترقيم الدولي الموحد للطباعة ISSN: 3009-612X
الترقيم الدولي الموحد الالكتروني E. ISSN: 3009-6146
المجلة معرفة على قاعدة المجلات العلمية فى بنك المعرفة المصرى
وقاعدة بيانات Edu Search دار المنظومة

Lawson, A. E. (2000). Classroom test of scientific reasoning. Revised Edition Journal of Research in Science Teaching, 15(1), 11-24.

Lawson, A. E. (2000). The generality of hypothetico-deductive reasoning. Science Education, 84(2), 183–194.

Lee, A. T., Hairston, R. V., Thames, R., Lawrence, T., & Herron, S. S. (2002). Using a computer simulation to teach science process skills to college biology and elementary majors. Bioscene, 28(4), 35.

Lehrer, R., & Schauble, L. (2000). Developing model-based reasoning in mathematics and science. Journal of Applied Developmental Psychology, 21(1), 39–48.

Liu, Q., Zhu, L., & Qin, Z. (2021). Using machine learning to predict student performance in science courses: A systematic approach. Journal of Science Education and Technology, 30(4), 532–545.

Liu, S., Zhu, M., & Qin, Z. (2021). Educational data mining for student performance prediction: A survey. IEEE Access, 9, 106954–106975.

Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. Advances in Neural Information Processing Systems, 30, 4765–4774.

Masnick, A. M., & Morris, B. J. (2008). The role of data characteristics in children's understanding of data. Child Development, 79(5), 1586–1601.

Next Generation Science Standards [NGSS]. (2013). Next generation science standards: For states, by states. The National Academies Press.

NGSS Lead States. (2013). Next Generation Science Standards: For states, by states. National Academies Press.

Nguyen, A., & Walker, M. (2016). A comparative review of pedagogical models for machine learning. In Proceedings of the 25th International Symposium on Machine Learning (pp. 32–40). ACM.

ISSN: 3009-612X الترقيم الدولي الموحد للطباعة
E. ISSN: 3009-6146 الترقيم الدولي الموحد الالكتروني
المجلة معرفة على قاعدة المجلات العلمية فى بنك المعرفة المصرى
وقاعدة بيانات Edu Search دار المنظومة

Osterhaus, C., Brandone, A. C., Vosniadou, S., & Nicolopoulou, A. (2021). The emergence and development of scientific thinking during the early years: Basic processes and supportive contexts. Frontiers in Psychology, 12, 629384.

Owusu Achiaw, A., & Owusu, K. A. (2023). Assessing Ghanaian primary school pupils' scientific reasoning skills. Eurasian Journal of Science and Environmental Education, 3(2), 99–107.

Padilla, M. J. (1990). The Science Process Skills. (Research Matters – to the Science Teacher No. 9004). Manhattan, KS: National Association for Research in Science Teaching.

Padilla, M. J. (1990). The Science Process Skills. In M. Keenan (Ed.), Research Matters - to the Science Teacher (No. 9004). National Association for Research in Science Teaching (NARST).

Pickal, A. J., Engelmann, K., Chinn, C. A., Neuhaus, B. J., Girwidz, R., & Wecker, C. (2023, June). The diagnosis of scientific reasoning skills: how teachers' professional knowledge predicts their diagnostic accuracy. In Frontiers in Education (Vol. 8, p. 1139176). Frontiers Media SA.

Pisa, O. E. C. D. (2019). Results (Volume I): What students know and can do. Organization for Economic Cooperation and Development (OECD): Paris, France.

Romero, C., & Ventura, S. (2020). Educational data mining and learning analytics: An updated survey. Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery, 10(3), e1355.

Scikit-learn Developers. (2023). RandomForestClassifier — scikit-learn 1.7.2 documentation. Retrieved from https://scikit-learn.org/stable/modules/generated/sklearn.ensemble.RandomForestClassifier.html

الترقيم الدولي الموحد للطباعة ISSN: 3009-612X
الترقيم الدولي الموحد الالكتروني E. ISSN: 3009-6146
المجلة معرفة على قاعدة المجلات العلمية فى بنك المعرفة المصرى
وقاعدة بيانات Edu Search دار المنظومة

Shute, V. J., & Ventura, M. (2013). Stealth Assessment: Measuring and Supporting Learning in Video Games. MIT Press.

Slavin, R. E. (2020). Educational psychology: Theory and practice (13th ed.). Pearson.

Sodian, B., Zaitchik, D., & Carey, S. (1991). Young children's differentiation of hypotheses and evidence. Child Development, 62(4), 753–766.

Tobin, K. G., & Capie, W. (1982). Development and validation of a group test of integrated science process skills. Journal of Research in Science Teaching, 19(2), 137–146.

Vygotsky, L. S. (1978). Mind in Society: The Development of Higher Psychological Processes. Cambridge, MA: Harvard University Press.

Woolf, B. P. (2010). Building Intelligent Interactive Tutors: Student-Centered Strategies for Revolutionizing E-Learning (2nd ed.). Morgan Kaufmann.

Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education–Are the technologies ready to impact practice? Computers & Education, 147, 103778.

Zimmerman, C. (2007). The development of scientific thinking skills in elementary and middle school. Developmental Review, 27, 172–223.

ISSN: 3009-612X الترقيم الدولي الموحد للطباعة
E. ISSN: 3009-6146 الترقيم الدولي الموحد الالكتروني
المجلة معرفة على قاعدة المجلات العلمية فى بنك المعرفة المصرى
وقاعدة بيانات Edu Search دار المنظومة