مجله مرکز مطالعات و توسعه آموزش علوم پزشکی یزد
دوره بیستم، شماره اول، بهار ۱۴۰۴
صفحه: ۱۱۴۵- ۱۱۳۵

Journal of Medical Education and Development
Vol. 20, No. 1, Spring 2025
Pages: 1135– 1145

گزارش کوتاه

ارزیابی پنهان در یادگیری مبتنی‌بر بازی: مطالعه مرور روایتی

نجمه زمانی[1]، حمید یزدانی‌نژاد[2,3*]

چکیده

مقدمه: یادگیری مبتنی‌بر بازی به عنوان یک روش آموزشی جذاب، پتانسیل بالایی در آموزش پزشکی دارد. ارزیابی پنهان که به طور یکپارچه در بازی ادغام می‌شود، می‌تواند اطلاعات ارزشمندی در مورد عملکرد فراگیران بدون ایجاد اضطراب ناشی از آزمون‌های سنتی ارائه دهد. هدف این مطالعه، تعیین کاربردها، مزایا و چالش‌های ارزیابی پنهان در یادگیری مبتنی بر بازی در آموزش پزشکی است.

روش بررسی: این مطالعه مروری روایتی با جستجوی نظام‌مند درپایگاه‌های داده PubMed، Scopus، Web of Science، ERIC، مگیران، ایران داک و SID در طی سال‌های ۲۰۱۰ تا ۲۰۲۵ انجام شد. کلمات کلیدی مورد استفاده در جستجو شامل "ارزیابی پنهان"، "آموزش"، "ارزیابی تکوینی" و "یادگیری مبتنی‌بر بازی" بودند. غربالگری، استخراج داده و تحلیل محتوایی به‌صورت مستقل توسط دو پژوهشگر و براساس معیارهای ورود صورت گرفت.

نتایج: در نهایت ۲۵ مطالعه مورد بررسی قرار گرفت. یافته‌ها نشان می‌دهند که ارزیابی پنهان می‌تواند بازخورد شخصی‌سازی شده، ارزیابی مهارت‌های غیرشناختی، افزایش انگیزه و ارائه داده‌های معتبر در یادگیری مبتنی‌بر بازی فراهم کند. همچنین نتایج نشان داد ارزشیابی پنهان با چالش‌هایی در زمینه‌های طراحی دقیق ارزیابی، اعتبار و پایایی داده‌ها، ملاحظات اخلاقی و هزینه و زمان روبروست.

نتیجه‌گیری: ارزیابی پنهان در محیط‌های یادگیری مبتنی‌بر بازی می‌تواند شیوه سنجش و یادگیری را متحول کند و فرصت‌های جدیدی برای بازخورد شخصی‌سازی‌شده، سنجش مهارت‌های پیچیده و ارتقای خودتنظیمی فراهم سازد. اما موفقیت عملی آن نیازمند اعتبارسنجی دقیق، توجه به ابعاد اخلاقی و توسعه پژوهش‌های طولی و تطبیقی است.

کلیدواژه‌ها: ارزیابی پنهان، یادگیری مبتنی‌بر بازی، آموزش، ارزیابی تکوینی

۱- دانشجوی کارشناسی ارشد پرستاری داخلی جراحی، دانشکده پرستاری مامایی، دانشگاه علوم پزشکی کاشان، ایران.
۲- مربی، گروه اتاق عمل و هوشبری، دانشکده پیراپزشکی، دانشگاه علوم پزشکی شهید صدوقی، یزد، ایران.
۳- مربی، مرکز مطالعات و توسعه آموزش علوم پزشکی، دانشگاه علوم پزشکی شهید صدوقی، یزد، ایران.
* (نویسنده مسئول)؛ تلفن: ۹۸۹۱۳۵۱۲۵۵۰۴+ پست الکترونیکی:*hyazdaninejhad@gmail.com*

تاریخ دریافت: ۲۲ / ۰۲ / ۱۴۰۳ تاریخ بازبینی: ۲۰ / ۰۳ / ۱۴۰۴ تاریخ پذیرش: ۲۵ / ۰۳ / ۱۴۰۴

CC BY-NC 4.0

[Downloaded from jmed.ssu.ac.ir on 2026-11-07]

[DOI: 10.18502/jmed.v20i1.18988]

مقدمه

تحولات سریع در فناوری و علوم پزشکی، آموزش پزشکی را با چالش‌های بنیادینی روبه‌رو ساخته است که ضرورت بازنگری در روش‌های آموزشی سنتی را برجسته می‌کند. بررسی‌ها نشان می‌دهد روش‌های معمول مبتنی‌بر سخنرانی و رویکردهای منفعل، نمی‌توانند نیازهای پیچیده یادگیری امروز دانشجویان پزشکی یا الزامات دنیای واقعیِ بالینی را به‌درستی برآورده نمایند؛ به‌طوری که مطالعه‌ای در سال ۲۰۱۱ بیانگر ناکارآمدی این شیوه‌ها در ارتقای سطح مهارت عملی و تفکر انتقادی فراگیران است (۱). در پاسخ به این چالش‌ها، یادگیری مبتنی‌بر بازی (GBL[۲]) مورد توجه گسترده قرار گرفته است.

GBL با فراهم آوردن محیط‌های شبیه‌سازی‌شده، تعاملی و جذاب، امکان تمرین مکرر، دریافت بازخورد و افزایش انگیزه یادگیرندگان را دارد. شواهد نشان داده‌اند که این روش منجر به بهبود معنادار یادگیری عمیق، توسعه مهارت‌های بالینی، و ارتقاء یادگیری خودتنظیمی در مقایسه با آموزش سنتی شده است (۲،۳). همچنین با فراهم‌سازی زمینه‌ای امن، دانشجویان می‌توانند با چالش‌ها و سناریوهای واقعی و پیچیده بالینی روبه‌رو شده و توانایی حل مسئله خود را تقویت کنند (۴).

ارزیابی یکی از اجزای اساسی فرایند یادگیری است. اهمیت آن، نه‌تنها در سنجش پیشرفت، بلکه در اصلاح مسیر و ارائه بازخورد به یادگیرنده نهفته است (۵). با وجود این، ارزیابی‌های رایج مانند آزمون‌های سنتی، اغلب با افزایش اضطراب، کاهش انگیزه و عدم سنجش دقیق عملکرد بالینی همراه هستند (۶). در این راستا، ارزیابی پنهان[۳] در محیط‌های مبتنی‌بر بازی، رویکردی نوین و مبتنی‌بر داده تلقی می‌شود که داده‌های مربوط به رفتار و عملکرد یادگیرنده را ضمن فرایند یادگیری و بدون ایجاد وقفه یا استرس جمع‌آوری می‌کند (۳،۷). برخلاف ارزیابی‌های آشکار و مبتنی‌بر آزمون، ارزیابی پنهان با بهره‌گیری از الگوریتم‌های پیشرفته هوش مصنوعی و داده‌کاوی، امکان سنجش مهارت‌های پیچیده شناختی، عاطفی و حرفه‌ای را فراهم ساخته است (۸،۹).

مرور مطالعات نشان می‌دهد که پیاده‌سازی موفق ارزیابی پنهان در محیط‌های GBL، می‌تواند با ارائه بازخورد فوری و تحلیل دقیق داده‌های تعامل، موجب بهبود خودتنظیمی و شخصی‌سازی فرآیند یادگیری شود (۲،۱۰). البته این رویکرد با چالش‌هایی نظیر پیچیدگی پیاده‌سازی فناوری‌های داده‌کاوی، حصول اطمینان از اعتبار ارزیابی، ملاحظات اخلاقی پیرامون حفظ حریم خصوصی و مدیریت حجم انبوه داده‌ها نیز روبه‌روست (۹،۱۱).

با توجه به شواهد موجود و اهمیت فزاینده طراحی رویکردهای نوآورانه در آموزش پزشکی، این مقاله مروری با هدف تبیین تاریخچه و روند توسعه ارزیابی پنهان در محیط‌های GBL، معرفی کاربردهای فعلی و مزایای اثبات‌شده آن، و واکاوی چالش‌ها و ملاحظات اخلاقی مرتبط تدوین شده است. پرسش‌های کلیدی مقاله به‌ترتیب عبارت‌اند از:

چگونه ارزیابی پنهان در محیط‌های مبتنی‌بر بازی در آموزش پزشکی پیاده‌سازی شده است؟

اصلی‌ترین مزایا و شواهد اثربخشی آن کدام‌اند؟

چه چالش‌های فناورانه، روان‌سنجی و اجرایی در پیاده‌سازی این رویکرد وجود دارد؟

و نهایتاً، چه ملاحظات اخلاقی و توصیه‌های عملی برای استفاده ایمن و مؤثر از ارزیابی پنهان در آموزش پزشکی باید مدنظر قرار گیرد؟

روش بررسی

مطالعه حاضر به صورت یک مطالعه مروری روایتی بود. معیارهای ورود به مطالعه شامل مقالات اصیل، مقالات مروری و گزارش‌های موردی به زبان فارسی و انگلیسی بودند که به طور مستقیم به موضوع ارزیابی پنهان در GBL در آموزش پرداخته بودند. مقالاتی که به طور مستقیم به موضوع مورد نظر مرتبط

[2] Game-based learning
[3] Stealthy Assessment

CC BY-NC 4.0

[Downloaded from jmed.ssu.ac.ir on 2026-11-07]

[DOI: 10.18502/jmed.v20i1.18988]

نبودند، از مطالعه حذف شدند. همچنین، مقالاتی که کیفیت روش‌شناختی پایینی داشتند نیز از مطالعه حذف شدند.

برای این منظور پایگاه‌های اطلاعاتی بین المللی مانند PubMed، Scopus، Web of Science و ERIC برای بازیابی مقالات انگلیسی و پایگاه های اطلاعاتی داخلی مگیران، ایران داک و SID برای بازیابی مقالات فارسی مورد جستجو قرار گرفتند. کلمات کلیدی مورد استفاده در جستجو شامل "stealthy assessment"، "education"، "formative assessment" و "game based learning" با معادل فارسی "ارزیابی پنهان"، "آموزش"، "ارزیابی تکوینی" و "یادگیری مبتنی‌بر بازی" بودند. برای افزایش دقت و جامعیت جستجو، از ترکیبات مختلف این کلمات کلیدی و مترادف‌های آنها نیز استفاده شد. مقالات منتشر شده در بازه زمانی ۲۰۱۰ تا ۲۰۲۵ مورد بررسی قرار گرفتند.

مطالعات بازیابی‌شده، براساس استراتژی جستجو وارد نرم افزار EndNote شدند. در ابتدا، مقالات تکراری با استفاده از نرم‌افزار شناسایی و حذف گردیدند. سپس عنوان و چکیده کلیه مطالعات بر اساس معیارهای ورود بررسی و در صورت نیاز متن کامل آن ها مطالعه شد. فرآیند انتخاب مطالعات توسط دو نفر از پژوهشگران به صورت مستقل انجام شد و در صورت هر گونه عدم توافق، به پژوهشگر سوم ارجاع گردید. پس از انتخاب مطالعات، طبق معیارهای ورود و خروج، گردآوری داده ها با استفاده از فرم استخراج داده ها که مبتنی‌بر اهداف مطالعه بود، انجام گردید. استخراج داده ها توسط دو نفر از اعضای تیم تحقیق به صورت مستقل انجام گردید. تجزیه و تحلیل داده ها از طریق روش آنالیز تحلیل محتوا انجام شد.

جدول ۱: خلاصه‌ای از استراتژی جستجو در پایگاه‌ها

Time limitation	2010 -2025
Language limitation	English AND Persian
Databases	Web of science, PubMed, Scopus, SID, ایران داک،مگیران
PubMed	("stealthy assessment"[Title/Abstract] OR "stealth assessment"[Title/Abstract]) AND ("education"[MeSH Terms] OR "education"[Title/Abstract]) AND ("formative assessment"[Title/Abstract] OR "formative evaluation"[Title/Abstract] OR "formative feedback"[Title/Abstract]) AND ("game based learning"[Title/Abstract] OR "game-based learning"[Title/Abstract] OR "gamification"[Title/Abstract])
Scopus	(TITLE-ABS-KEY("stealthy assessment" OR "stealth assessment")) AND (TITLE-ABS-KEY("education")) AND (TITLE-ABS-KEY("formative assessment" OR "formative evaluation" OR "formative feedback")) AND (TITLE-ABS-KEY("game based learning" OR "game-based learning" OR "gamification"))
Web of Science	TS=("stealthy assessment" OR "stealth assessment") AND TS=("education") AND TS=("formative assessment" OR "formative evaluation" OR "formative feedback") AND TS=("game based learning" OR "game-based learning" OR "gamification")

یافته‌ها

در نهایت ۲۵ مقاله وارد مطالعه گردید. بررسی ها نشان داد که ارزیابی پنهان می‌تواند فرصت‌های متعددی را برای بهبود یادگیری و ارزیابی در GBL فراهم کند. در این بخش، به تفصیل به بررسی این فرصت‌ها و چالش‌های مربوط به آنها می‌پردازیم:

CC BY-NC 4.0

[Downloaded from jmed.ssu.ac.ir on 2026-11-07]

[DOI: 10.18502/jmed.v20i1.18988]

۱. ارائه بازخورد شخصی‌سازی شده

یکی از مهم‌ترین مزایای ارزیابی پنهان، امکان ارائه بازخورد شخصی‌سازی‌شده به فراگیران است. برخلاف رویکردهای سنتی که معمولاً بازخورد را پس از پایان آزمون و به شکل عمومی ارائه می‌دهند، ارزیابی پنهان می‌تواند بر اساس داده‌های رفتاریِ فردی، بازخوردی هدفمند و به‌هنگام فراهم کند. این ویژگی بر مبنای مولفه «بازخورد فوری و معنادار» در نظریه بار شناختی[۴] است که تأکید دارد بار اضافی ذهنی باید با بازخورد مناسب کاهش یابد تا یادگیری عمیق‌تر اتفاق بیافتد. این نوع بازخورد می‌تواند به فراگیر کمک کند تا اشتباهات خود را شناسایی کرده و در آینده از تکرار آنها جلوگیری کند.

چگونگی ارائه بازخورد شخصی‌سازی شده:

- شناسایی نقاط قوت و ضعف: با تجزیه و تحلیل داده‌های جمع‌آوری شده از بازی، می‌توان نقاط قوت و ضعف فراگیران را شناسایی کرد. به عنوان مثال، اگر فراگیری در تشخیص یک بیماری خاص مشکل دارد، می‌توان بازخورد خاصی را در مورد آن بیماری به او ارائه داد.
- ارائه بازخورد فوری: بازخورد باید به صورت فوری و در زمان مناسب ارائه شود. به عنوان مثال، اگر فراگیری در حین بازی یک اشتباه مرتکب می‌شود، می‌توان بلافاصله بازخورد اصلاحی را به او ارائه داد.
- ارائه بازخورد نقادانه: بازخورد باید نقادانه باشد و فراگیر را به فکر کردن و تحلیل عملکرد خود تشویق کند. به عنوان مثال، می‌توان از فراگیر خواست که توضیح دهد چرا یک تصمیم خاص را گرفته است و چه چیزهایی را در نظر داشته است.

مثال عملی:

در یک بازی شبیه‌سازی بیمار، فراگیر نقش یک پزشک را بازی می‌کند و باید بیماران مختلف را تشخیص دهد و درمان کند. ارزیابی پنهان می‌تواند با جمع‌آوری داده‌های مربوط به تصمیمات فراگیر در طول بازی، بازخورد شخصی‌سازی شده‌ای را به او ارائه دهد. به عنوان مثال، اگر فراگیری در تشخیص یک بیماری خاص اشتباه می‌کند، بازی می‌تواند به او بازخورد زیر را ارائه دهد:

"شما در تشخیص بیماری X اشتباه کردید. به نظر می‌رسد که شما علائم Y و Z را نادیده گرفتید. لطفاً دوباره به این علائم توجه کنید و سعی کنید تشخیص خود را اصلاح کنید." (۳،۲).

۲. ارزیابی مهارت‌های غیرشناختی

برخلاف آزمون‌های معمول، ارزیابی پنهان این توان را دارد تا ابعاد پیچیده‌تری چون مهارت‌های غیرشناختی (تفکر انتقادی، کار تیمی، حل مسئله و...) را نیز پوشش دهد. برای این منظور، لازم است شاخص‌های رفتاری عینی و قابل سنجش تعریف گردد. شاخص‌های رفتاری عبارتند از:

- برای ارزیابی «تفکر انتقادی»: تعداد دفعاتی که فراگیر چند گزینه را قبل از انتخاب نهایی بررسی می‌کند، یا میزان استناد او به داده‌های بالینی برای تصمیم‌گیری.
- برای «کار تیمی»: دفعات مشارکت در حل مسئله گروهی، درصد تعاملات مؤثر با سایر اعضا، یا میزان کمک به هم‌تیمی‌ها هنگام مواجهه با یک چالش.

چگونگی ارزیابی مهارت‌های غیرشناختی:

- تعریف شاخص‌های رفتاری: برای ارزیابی هر یک از مهارت‌های غیرشناختی، باید شاخص‌های رفتاری مشخصی تعریف کرد. به عنوان مثال، برای ارزیابی تفکر انتقادی، می‌توان شاخص‌هایی مانند توانایی تحلیل اطلاعات، شناسایی فرضیات و ارزیابی شواهد را در نظر گرفت.
- جمع‌آوری داده‌های رفتاری: داده‌های مربوط به رفتار فراگیران در بازی باید به طور مداوم جمع‌آوری شوند. این داده‌ها می‌تواند شامل تصمیماتی باشد که فراگیر در طول بازی می‌گیرد، تعاملاتی که با سایر بازیکنان دارد و نحوه حل مسائلی که با آنها روبرو می‌شود.

[4] Cognitive Load Theory

CC BY-NC 4.0

[Downloaded from jmed.ssu.ac.ir on 2026-11-07]

[DOI: 10.18502/jmed.v20i1.18988]

• تجزیه و تحلیل داده‌های رفتاری: داده‌های جمع‌آوری شده باید با استفاده از روش‌های مناسب تجزیه و تحلیل شوند. به عنوان مثال، می‌توان از روش‌های تحلیل محتوا برای بررسی نحوه استدلال فراگیران در طول بازی استفاده کرد.

مثال عملی:

در یک بازی شبیه‌سازی یک تیم پزشکی، فراگیر باید با سایر بازیکنان همکاری کند تا بیماران را درمان کند. ارزیابی پنهان می‌تواند با جمع‌آوری داده‌های مربوط به تعاملات فراگیر با سایر بازیکنان، مهارت‌های کار تیمی او را ارزیابی کند. به عنوان مثال، بازی می‌تواند میزان مشارکت فراگیر در بحث‌ها، توانایی او در گوش دادن به نظرات دیگران و تمایل او به همکاری با سایر بازیکنان را ارزیابی کند (۶،۷).

۳. افزایش انگیزه و تعامل

ارزیابی پنهان می‌تواند با ادغام ارزیابی به طور یکپارچه در بازی، انگیزه و تعامل فراگیران را افزایش دهد. در ارزیابی‌های سنتی، فراگیران اغلب احساس می‌کنند که در حال امتحان دادن هستند و این موضوع می‌تواند باعث اضطراب و استرس آنها شود. در مقابل، در ارزیابی پنهان، فراگیران ممکن است حتی متوجه نشوند که در حال ارزیابی شدن هستند، که می‌تواند اضطراب و استرس آنها را کاهش دهد. همچنین در ادبیات GBL، ادغام ارزیابی در بازی موجب افزایش انگیزه و مشارکت فعال می‌شود. بر اساس نظریه خودتعیینی، انگیزش ذاتی زمانی بیشتر می‌شود که افراد بتوانند حس شایستگی، خودمختاری و ارتباط مؤثر با دیگران را تجربه کنند.

چگونگی افزایش انگیزه و تعامل:

• جذاب و سرگرم‌کننده بودن بازی: بازی باید جذاب و سرگرم‌کننده باشد تا فراگیران از انجام آن لذت ببرند.

• ارائه چالش‌های مناسب: چالش‌های ارائه شده در بازی باید مناسب سطح دانش و مهارت فراگیران باشند. چالش‌های خیلی آسان باعث بی‌حوصلگی فراگیران می‌شوند و چالش‌های خیلی سخت باعث ناامیدی آنها می‌شوند.

• ارائه بازخورد مثبت: بازخورد مثبت می‌تواند به فراگیران انگیزه دهد تا به تلاش خود ادامه دهند.

مثال عملی:

در یک بازی آموزشی در مورد آناتومی بدن انسان، فراگیر باید قطعات مختلف بدن را در جای مناسب قرار دهد. ارزیابی پنهان می‌تواند با ارائه امتیاز و پاداش به فراگیر در ازای هر قطعه‌ای که به درستی در جای خود قرار می‌دهد، انگیزه او را افزایش دهد (۱،۱۱).

۴. ارائه داده‌های معتبر و قابل اعتماد

ارزیابی پنهان می‌تواند داده‌های معتبر و قابل اعتمادی را در مورد عملکرد فراگیران ارائه دهد. در ارزیابی‌های سنتی، داده‌ها معمولاً در یک مقطع زمانی خاص جمع‌آوری می‌شوند و ممکن است تصویر دقیقی از عملکرد فراگیر در طول زمان ارائه ندهند. در مقابل، ارزیابی پنهان می‌تواند داده‌ها را به طور مداوم و در طول زمان جمع‌آوری کند، که می‌تواند به ارائه یک تصویر جامع از پیشرفت فراگیر کمک کند.

چگونگی ارائه داده‌های معتبر و قابل اعتماد:

• استفاده از روش‌های معتبر ارزیابی: برای جمع‌آوری داده‌ها باید از روش‌های معتبر ارزیابی استفاده شود.

• جمع‌آوری داده‌ها به طور مداوم: داده‌ها باید به طور مداوم و در طول زمان جمع‌آوری شوند.

• تجزیه و تحلیل داده‌ها با استفاده از روش‌های آماری: داده‌ها باید با استفاده از روش‌های آماری مناسب تجزیه و تحلیل شوند.

مثال عملی:

در یک بازی شبیه‌سازی جراحی، فراگیر باید مراحل مختلف جراحی را انجام دهد. ارزیابی پنهان می‌تواند با جمع‌آوری داده‌های مربوط به دقت و سرعت فراگیر در انجام هر یک از مراحل جراحی، داده‌های معتبری را در مورد عملکرد او ارائه دهد (۱۰،۹).

[Downloaded from jmed.ssu.ac.ir on 2026-11-07]

[DOI: 10.18502/jmed.v20i1.18988]

CC BY-NC 4.0

جدول ۲: فرصت‌های کلیدی ارزیابی پنهان در GBL

فرصت اصلی	شرح مختصر	مثال عملی	منابع
بازخورد شخصی‌سازی‌شده	ارائه بازخورد هدفمند و به‌هنگام بر اساس داده‌های رفتاری فردی (کاهش بار شناختی و ارتقای یادگیری عمیق)	بازی شبیه‌سازی بالینی: هنگام تشخیص نادرست بیماری X، نمایش پیام «علائم Y و Z را مجدداً بررسی کنید»	۲،۳
ارزیابی مهارت‌های غیرشناختی	استخراج شاخص‌های رفتاری برای سنجش تفکر انتقادی (مثلاً دفعات بررسی گزینه‌ها) و کار تیمی (مثلاً میزان مشارکت در تعامل گروهی)	بازی تیمی پزشکی: اندازه‌گیری میزان مشارکت در تصمیم‌گیری مشترک و کمک به هم‌تیمی‌ها	۶،۷
افزایش انگیزه و مشارکت فعال	ادغام ارزیابی در فرایند بازی بدون القای حس «امتحان» >> کاهش اضطراب و تقویت انگیزه ذاتی و احساس خودمختاری و شایستگی	بازی آناتومی: امتیازدهی و پاداش آنی به فراگیر برای قرار دادن صحیح قطعات بدن	۱،۱۱
تولید داده‌های معتبر و قابل‌اعتماد	ثبت پیوسته داده‌ها در طول بازی (سرعت و دقت تصمیم‌گیری در هر مرحله) >> تصویر جامع از پیشرفت یادگیرنده	شبیه‌سازی جراحی: ضبط زمان و دقت هر مرحله و تحلیل آماری برای ارزیابی پیشرونده مهارت‌های جراحی	۹،۱۰

چالش‌های ارزیابی پنهان در GBL

در کنار مزایای فراوانی که ارزیابی پنهان در GBL دارد، چالش‌هایی نیز در طراحی و پیاده‌سازی این نوع ارزیابی وجود دارد که باید به آنها توجه شود:

۱. طراحی دقیق ارزیابی: طراحی ارزیابی پنهان نیازمند دقت و تخصص بالایی است. ارزیابی باید به گونه‌ای طراحی شود که بتواند اطلاعات معتبر و قابل اعتمادی را در مورد دانش، مهارت‌ها و رفتار یادگیرندگان جمع‌آوری کند. بنابراین باید شاخص‌های ارزیابی به طور واضح و دقیق تعریف شوند و روش‌های جمع‌آوری داده‌ها به گونه‌ای طراحی شوند که بتوانند این شاخص‌ها را به طور موثر اندازه‌گیری کنند (۱).

۲.اعتبار و پایایی داده‌ها: اطمینان از اعتبار و پایایی داده‌های جمع‌آوری شده از طریق ارزیابی پنهان یک چالش مهم است. باید از روش‌های مناسب برای تجزیه و تحلیل داده‌ها استفاده شود تا از تفسیر صحیح نتایج اطمینان حاصل شود. در همین راستا باید از روش‌های آماری مناسب برای بررسی اعتبار و پایایی داده‌ها استفاده شود و در صورت لزوم، داده‌ها اصلاح شوند (۹).

۳. ملاحظات اخلاقی: استفاده از ارزیابی پنهان ملاحظات اخلاقی خاصی را به همراه دارد. یادگیرندگان باید از جمع‌آوری داده‌ها و نحوه استفاده از آنها آگاه باشند. همچنین، باید از حفظ حریم خصوصی و محرمانه بودن اطلاعات یادگیرندگان اطمینان حاصل شود. برای این منظور، باید از فراگیران رضایت آگاهانه گرفته شود و سیاست‌های حفظ حریم خصوصی به طور واضح تعریف شوند (۱۱).

۴.هزینه و زمان: طراحی و توسعه GBL با ارزیابی پنهان ممکن است پرهزینه و زمان‌بر باشد. نیاز به تخصص در زمینه‌های مختلف مانند طراحی بازی، آموزش پزشکی و ارزیابی می‌تواند هزینه‌ها را افزایش دهد. برای کاهش هزینه‌ها و زمان، می‌توان از ابزارها و فناوری‌های موجود استفاده کرد و از همکاری با متخصصان مختلف بهره برد (۲).

[Downloaded from jmed.ssu.ac.ir on 2026-11-07]

[DOI: 10.18502/jmed.v20i1.18988]

CC BY-NC 4.0

جدول ۳: چالش‌ها و ملاحظات طراحی و پیاده‌سازی ارزیابی پنهان

چالش	شرح مختصر	پیشنهاد راهکار	منابع
طراحی دقیق شاخص‌های ارزیابی	نیاز به تعریف واضح و عملیاتی‌سازی معیارهای رفتاری برای دانش، مهارت و رفتار یادگیرندگان	استفاده از روش دلفی و تحلیل محتوا با مشارکت تیمی میان‌رشته‌ای برای تعیین و اجماع بر شاخص‌های کلیدی	۱
اعتبار و پایایی داده‌ها	اطمینان از صحت و ثبات داده‌های جمع‌آوری‌شده (خطر سوگیری در ثبت و تفسیر داده‌های بازی)	بهره‌گیری از مدل‌های آماری (تحلیل عاملی، آزمون–بازآزمون) و استانداردسازی فرایند جمع‌آوری و پاک‌سازی داده‌ها	۹
ملاحظات اخلاقی	شفافیت در نحوه ثبت و استفاده از داده‌ها، رضایت آگاهانه کاربران و حفظ محرمانگی اطلاعات	تدوین چارچوب حاکمیت داده (Data Governance)، فرایند اخذ رضایت آگاهانه و اطلاع‌رسانی شفاف به یادگیرندگان	۱۱
هزینه و زمان توسعه	نیاز به تخصص‌های طراحی بازی، روان‌سنجی و علوم داده >> افزایش هزینه و طولانی شدن چرخه توسعه	استفاده از پلتفرم‌های آماده و ابزارهای متن‌باز، و ایجاد همکاری بین دانشگاه‌ها و صنعت برای اشتراک منابع و کاهش هزینه‌ها	۲

بحث

نتایج این مطالعه مروری نشان می‌دهد که ارزیابی پنهان در محیط‌های بازی‌محور، پتانسیل بالایی برای بهبود انگیزش، خودتنظیمی و مشارکت فعال یادگیرندگان دارد. مطابق مطالعات اولیه، جمع‌آوری بی‌درنگ داده‌های رفتاری مانند کلیک‌ها، زمان واکنش و مسیرهای کاربری می‌تواند محرکی قدرتمند برای ارتقای خودکارآمدی و درگیری شناختی فراگیران باشد (۱،۲،۶). این نتایج با اصول نظریه خودتعیینی و مفهوم «جریان[5]» هم‌راستا است و نشان می‌دهد ادغام بهینه چالش‌ها در بازی می‌تواند محیطی فراهم آورد که نیازهای اساسی روانی یادگیرندگان را تأمین کند و ورود به حالت جریان را تسهیل سازد (۱۴،۱۵).

از منظر روان‌سنجی، مطالعات متعددی با به‌کارگیری مدل‌های یادگیری عمیق مانند شبکه‌های عصبی عمیق (۳،۸) و LSTM[6] (۴) سعی در اعتبارسنجی ابزارهای ارزیابی پنهان داشته‌اند. تداخل نتایج بازی‌های رایانه‌ای با آزمون‌های استاندارد و تحلیل عاملی بارها گزارش شده، اما هنوز سلسله‌مراتبی رسمی برای استانداردسازی این فرایندها تدوین نشده است (۲۳،۴،۹،۱۰،۲). بنابراین، تشکیل کنسرسیومی از مراکز تحقیقاتی و روان‌سنجان برای ایجاد دستورالعمل‌های اعتبارسنجی و پایایی بلندمدت بسیار ضروری است.

تعریف شاخص‌های رفتاری مورد سنجش یکی دیگر از چالش‌های کلیدی است. اگرچه مدل‌های LSTM و تحلیل‌های چندمتغیره امکان پیش‌بینی مهارت‌هایی مانند تفکر انتقادی، خلاقیت و کار تیمی را فراهم آورده‌اند (۳،۴،۷،۲۰،۲۱،۲۶)، نبود توافق بر سر متغیرهای کلیدی منجر به برداشت‌های ناهمگون و خطرات تفسیر اشتباه داده می‌شود. پیشنهاد می‌شود پیش از مرحله پیاده‌سازی، با روش‌های ساخت‌یافته‌ای نظیر دلفی یا تحلیل محتوا با مشارکت خبرگان، شاخص‌های رفتاری مشخص و مستندسازی شوند.

طراحی بازی نیز باید مدیریت دقیقی بر بار شناختی یادگیرندگان داشته باشد. استفاده از سازوکارهای تطبیقی بر مبنای Adaptive Hypermedia (۱۲) و بهره‌گیری از گزارش‌های بلادرنگ برای تنظیم خودکار سطح دشواری می‌تواند از کاهش انگیزش یا افزایش بار شناختی ناخواسته جلوگیری کند (۲۹،۳۰). افزون بر این، به‌کارگیری الگوریتم‌های

[5] Flow
[6] long short-term memory

CC BY-NC 4.0

[Downloaded from jmed.ssu.ac.ir on 2026-11-07]

[DOI: 10.18502/jmed.v20i1.18988]

یادگیری ماشینی برای تنظیم پویا چالش‌ها، کلیدی است تا همواره تعادل میان سطح مهارت و دشواری حفظ شود.

در زمینه فرایند اجرا، همکاری میان‌رشته‌ای میان روان‌سنج‌ها، طراحان بازی، متخصصان علوم داده و آموزش پزشکی امری بدیهی است (۲،۶،۲۵). کاهش دغدغه‌های اخلاقی شامل شفافیت فرایند ارزیابی، جلب رضایت آگاهانه و تضمین حریم خصوصی داده‌ها، مستلزم تدوین چارچوب‌های حکمرانی داده از ابتداست (۱۸،۳۰). همچنین، ارزیابی هزینه-فایده در پیاده‌سازی کلینیکی یا آموزشی می‌تواند تصمیم‌گیری ذی‌نفعان را تسهیل کند (۲۸).

در افق پژوهش‌های آینده، باید به سوی مطالعات چندمرکزی، با حجم نمونه بالا و طرح‌های تصادفی کنترل‌شده یا پیش‌آزمون—پس‌آزمون حرکت کرد تا اعتبار نتایج افزایش یابد و مقایسه‌ای عینی با روش‌های سنتی فراهم شود (۹،۱۰،۱۹،۲۹). توسعه داشبوردهای تحلیلی و ابزارهای پیشرفته تجسم داده، امکان خودتنظیمی یادگیرنده و کاوش عمیق در داده‌های آموزشی را افزایش می‌دهد و مسیر تحقیق در این حوزه را هموار می‌سازد (۶،۲۱).

ارزیابی پنهان در بازی‌های آموزشی می‌تواند زمینه‌ساز شخصی‌سازی یادگیری، اندازه‌گیری مهارت‌های پیچیده و تقویت انگیزش باشد. برای بهره‌برداری کامل از این روش، مستلزم تدوین معیارهای روان‌سنجی قوی، استانداردسازی شاخص‌ها، توسعه فناوری‌های تطبیقی، رعایت ملاحظات اخلاقی و طراحی مطالعات چندمرکزی سخت‌گیرانه هستیم.

نتیجه گیری

ارزیابی پنهان به عنوان یک رویکرد نوین و داده‌محور در محیط‌های یادگیری مبتنی‌بر بازی، می‌تواند با ارائه بازخورد فوری، معتبر و شخصی‌سازی‌شده، تحول اساسی در سنجش و ارتقاء کیفیت یادگیری در آموزش پزشکی ایجاد کند. این رویکرد از سویی امکان سنجش مؤلفه‌های پیچیده‌ای مانند مهارت‌های غیرشناختی، انگیزه و تعامل اجتماعی را فراهم می‌آورد و از سوی دیگر، زمینه بهبود خودتنظیمی و شخصی‌سازی فرایند آموزش را مهیا می‌سازد.

با این حال، چالش‌هایی جدی از جمله اعتبارسنجی شاخص‌های رفتاری، تضمین منصفانه بودن سنجش، پیچیدگی فنی، مسائل اخلاقی و هزینه‌های پیاده‌سازی، محدودیت‌هایی پیش روی توسعه عملیاتی این رویکرد ایجاد کرده‌اند. به ویژه، خلأ مطالعات طولی و تجربی در زمینه پیامدهای پایدار و مقایسه با سنجش‌های سنتی همچنان وجود دارد.

بر این اساس، جهت بهره‌برداری بهینه و ایمن از ارزیابی پنهان در آموزش پزشکی، پیشنهاد می گردد:

- طراحی ابزارهای ارزیابی مبتنی‌بر شواهد روان‌سنجی و نظریه‌های یادگیری انجام شود؛
- فرآیند اعتبارسنجی و پایایی به صورت مستمر و تطبیقی دنبال گردد؛
- فاکتورهای اخلاقی از جمله شفافیت، آگاهی یادگیرندگان و حفظ حریم خصوصی به طور جدی رعایت شود؛
- و در نهایت، پژوهش‌های آتی به سمت مطالعات کارآزمایی تصادفی و مقایسه‌ای، تحلیل تأثیر طولی و بررسی اثربخشی در بسترها و جمعیت‌های متنوع گسترش یابد.

ارزیابی پنهان، با ادغام مفاهیم فناورانه، آموزشی و روان‌سنجی، می‌تواند به عنوان یکی از مؤثرترین ابزارهای تحول آموزش پزشکی در دهه آینده معرفی شود، مشروط به آنکه ابعاد فناورانه، اخلاقی و شواهد تجربی به‌درستی مدیریت و توسعه یابند.

مشارکت نویسندگان

ن، زمانی: تحلیل و تفسیر داده‌ها، نگارش پیش نویس اولیه مقاله

ح یزدانی نژاد: طراحی مطالعه، بازبینی انتقادی و اصلاحات علمی، تأیید نهایی نسخه ارسالی، پاسخگویی به سؤالات و ارتباط مکاتباتی

CC BY-NC 4.0

[Downloaded from jmed.ssu.ac.ir on 2026-11-07]

[DOI: 10.18502/jmed.v20i1.18988]

تعارض و منافع

نویسندگان اعلام می‌دارند تعارض منابفعی نداشته اند.

ملاحظات اخلاقی

نویسندگان تمام اصول اخلاق رعایت گردیده است و این مطالعه مروری می باشد.

حمایت مالی

این مطالعه از سوی هیچ سازمانی حمایت مالی دریافت نکرده است.

References

1. Shute VJ. ***Stealth assessment in computer-based games to support learning***. Computer games and instruction. 2011; 55(2): 503-24.
2. Sabourin JL. ***Stealth assessment of self-regulated learning in game-based learning environments***. North Carolina State University; 2013.
3. Min W, Frankosky MH, Mott BW, Rowe JP, Wiebe E, Boyer KE, Lester JC. ***DeepStealth: Leveraging deep learning models for stealth assessment in game-based learning environments.*** In Artificial Intelligence in Education: 17th Int Conference, AIED 2015, Madrid, Spain, June 22-26, 2015. Proceedings 17 2015 (pp. 277-286). Springer Int Publishing.
4. Akram B, Min W, Wiebe E, Mott B, Boyer KE, Lester J. ***Improving stealth assessment in game-based learning with LSTM-based analytics.*** In Int Conference on Educational Data Mining 2018 Jul.
5. Pool J. ***Stealth assessment and computer game learning: could this benefit the children of tomorrow?*** (Doctoral dissertation, Anglia Ruskin Research Online (ARRO)).
6. Srinivasa KG, Kurni M, Saritha K. ***Stealth Assessment. In Learning, Teaching, and Assessment Methods for Contemporary Learners: Pedagogy for the Digital Generation***. 2022; 263-281. Singapore: Springer Nature Singapore.
7. Ren X. ***Stealth assessment embedded in Game-Based Learning to measure soft skills: A critical review***. Game-based assessment revisited. 2019; 67-83.
8. Min W, Frankosky MH, Mott BW, Rowe JP, Smith A, Wiebe E, Boyer KE, Lester JC. ***DeepStealth: Game-based learning stealth assessment with deep neural networks. IEEE Transactions on Learning Technologies***. 2019; 13(2): 312-25.
9. Lu W, Griffin J, Sadler TD, Laffey J, Goggins SP. ***Game-Based Learning Prediction Model Construction: Toward Validated Stealth Assessment Implementation***. J of Learning Analytics. 2025; 12(1): 293-321.
10. Rajendran D, Prasanna S. ***Stealth assessment strategy in distributed systems using optimal deep learning with game-based learning***. The J of Supercomputing. 2022; 78(6): 8285-301.
11. Rahimi S, Shute VJ. ***Stealth assessment: a theoretically grounded and psychometrically sound method to assess, support, and investigate learning in technology-rich environments.*** Educational technology research and development. 2024; 72(5): 2417.
12. Brusilovsky P. ***Adaptive hypermedia. User Modeling and User-Adapted Interaction.*** 2001; 11(1-2): 87-110.
13. de-Juan-Ripoll C, Soler Domínguez JL, Chicchi Giglioli IA, Contero M, Alcañiz M. ***Virtual serious games for neuro-rehabilitation: A review of features and needs.*** J Med Syst. 2020; 44(8): 144.
14. Deci EL, Ryan RM. ***Self-determination theory in health care and its relations to motivational interviewing: A few comments.*** Int J Behav Nutr Phys Act. 2012; 9: 24.
15. Csikszentmihalyi M. ***Flow: The psychology of optimal experience.*** New York: Harper & Row; 1990.
16. DeRosier ME, CraigAshley B, Sanchez RP. ***Enhancing academic achievement in children with social and behavioral problems.*** In: Taras H, Potts-Datema W, editors. School Health: Policy and Practice. New York: Springer; 2012; 235-48.
17. Almond R, Tingir S, Lu X, Sun C, Rahimi S. ***A comparison of stealth assessment and traditional assessment in a digital reading environment.*** Computers in Human Behavior. 2017; 76:72-78.
18. Antoniou PE, Siountas A, Zilidou VI, Bamidis PD. ***Serious games as a tool for the management of mental health, stress and anxiety.*** Stud Health Technol Inform. 2017; 237: 60-67.

CC BY-NC 4.0

[Downloaded from jmed.ssu.ac.ir on 2026-11-07]

[DOI: 10.18502/jmed.v20i1.18988]

19. Bainbridge K, Shute VJ, Rahimi S, Liu X, Slater S, Baker RS, D'Mello S. ***Using stealth assessments to predict and improve learners' problem-solving skills with game-based learning.*** Computers in Human Behavior. 2022; 132: 107252.
20. DiCerbo KE. ***Game-based assessment of persistence***. Educ Psychol. 2014; 49(4): 243-56.
21. DiCerbo KE, Shute V, Kim YJ. ***Assessment through digital games.*** In: Zumbo BD, Hubley AM, editors. Understanding and Investigating Response Processes in Validation Research. Cham: Springer; 2016; 103-119.
22. Glaveanu VP, Hanchett Hanson M, Baer J, Barbot B, Clapp E, Hennessey B, et al. ***Advancing Creativity Theory and Research: A Socio-cultural Manifesto***. J Creative Behav. 2019; 53(3): 251–255.
23. NEVES-PEREIRA, Mônica Souza. **Creativity, Culture, and Education.** The Oxford Handbook of Creativity and Education, 2025, 379.
24. DiCerbo KE, Xu Y, Levy R, Lai E, Holland L. ***Automated scoring in game based assessment of persistence***. J Comput Assist Learn. 2017; 33(3): 263–277.
25. Al-Smadi M, Capuano N, Guetl C. ***Decoupling assessment and serious games to support guided exploratory learning in smart education.*** J of Ambient Intelligence and Humanized Computing. 2018; 9: 497-511.
26. Shute VJ and Rahimi S. ***Stealth assessment of creativity in a physics video game.*** Computers in Human Behavior. 2021; 116: 106647.
27. Dorans NJ. ***Contributions to the quantitative assessment of item, test, and score fairness. Advancing human assessment: The methodological, psychological and policy contributions of ETS***. 2017: 201-30.
28. Zhu S, Guo Q, Yang HH. ***Digital Game-Based Assessment on Student Evaluation: A Systematic Review.*** Int Conference on Blended Learning, Springer. 2022; 85-96.
29. Cham.Su F, Zou D. ***A Systematic Review of Game-Based Assessment in Education in the Past Decade.*** Knowledge Management & E-Learning. 2024; 16(3): 451-76.
30. Shute V, Ventura M. ***Stealth assessment: Measuring and supporting learning in video games.*** The mit press; 2013.

[Downloaded from jmed.ssu.ac.ir on 2026-11-07]

[DOI: 10.18502/jmed.v20i1.18988]

CC BY-NC 4.0

Stealthy Assessment in Game-Based Learning: A Narrative Review Study

Zamani N (MSc Student)[1], Yazdaninejad H(MSc)[2,3]*

[1] MSc Nursing Student, school of Nursing and Midwifery, Kashan University of Medical Sciences, Kashan, Iran
[2] Instructor, Department of Operating Room and Anesthesiology, School of Paramedical Sciences, Shahid Sadoughi University of Medical Sciences, Yazd, Iran
[3] Instructor, Educational Development Center, Shahid Sadoughi University of Medical Sciences, Yazd, Iran

Received: 12 May 2025 ***Revised:*** 10 June 2025 ***Accepted:*** 15 June 2025

Abstract

Introduction: Game-based learning (GBL) has great potential as an engaging learning method in medical education. Stealthy assessment, seamlessly integrated into the game, can provide valuable information about learners' performance without causing the anxiety caused by traditional tests. The aim of this study is to determine the applications, benefits, and challenges of Stealthy assessment in GBL in medical education.

Methods*:* The presented narrative review was conducted by systematically searching PubMed, Scopus, Web of Science, ERIC, MagIran, IranDoc, and SID databases between 2010 and 2025. The keywords used in the search included "Stealthy assessment", "education", "formative assessment," and "game-based learning". Screening, data extraction, and content analysis were performed independently by two researchers based on the inclusion criteria.

***Results*:** Ultimately, 25 studies were reviewed. Findings indicate that Stealthy assessment can provide personalized feedback, assessment of non-cognitive skills, increase in motivation, and valid data exhibition in game-based learning. Findings also indicate that Stealthy assessment faces challenges in the areas of accurate assessment design, data validity and reliability, ethical considerations, and cost and time.

Conclusion: Stealthy assessment in GBL environments can evolve the way we assess and learn, providing new opportunities for personalized feedback, measuring complex skills, and promoting self-regulation. But its practical success requires careful validation, consideration of ethical dimensions, and the development of longitudinal and comparative studies.

Keywords: stealthy assessment, game-based learning, education, formative assessment

This paper should be cited as:

Zamani N, Yazdaninejad H. ***Stealthy Assessment in Game-Based Learning: A Narrative Review Study.*** J Med Edu Dev 2025; 20(1): 1135- 1145.

**** Corresponding Author: Tel: +989135125504, Email: hyazdaninejhad@gmail.com***

[Downloaded from jmed.ssu.ac.ir on 2026-11-07]

[DOI: 10.18502/jmed.v20i1.18988]

CC BY-NC 4.0