

Avaliação da aprendizagem com o uso do ChatGPT: aprendizagens para a docência em química

Assessment of learning using ChatGPT: lessons for teaching chemistry

Evaluación del aprendizaje mediante ChatGPT: lecciones para la enseñanza de la química

Andréia Francisco Afonso¹
Augusto Theodoro de Carvalho²

Resumo

Este trabalho é resultado de uma pesquisa que teve como objetivo, investigar, na concepção dos licenciandos em Química, as contribuições e limitações do ChatGPT na elaboração de uma questão, para avaliar a aprendizagem de um conteúdo químico, abordado no Ensino Médio. A avaliação da aprendizagem é uma das atividades docentes e como pode ser elaborada com o auxílio da inteligência artificial generativa, é importante que os licenciandos reflitam sobre sua utilização. Para alcançarmos tal objetivo, elaboramos uma questão avaliativa, no âmbito da disciplina Avaliação do Processo de Ensino e Aprendizagem de Química, que está presente na matriz curricular das licenciaturas em Química de uma universidade federal, localizada em Minas Gerais, alinhando a temática avaliação junto ao uso da inteligência artificial generativa. A partir das respostas dos cinco licenciandos em Química, matriculados na disciplina, construímos os dados e os interpretamos pela Análise de Conteúdo. Verificamos que os licenciandos identificaram que o ChatGPT forneceu respostas que precisam ser revistas e corrigidas. Isso indica que os futuros professores não utilizarão as informações obtidas sem se certificarem que estão corretas, o que indica ser uma aprendizagem para a docência. Diante dos resultados obtidos, ressaltamos a importância da realização de outras pesquisas, pois ainda há uma lacuna na produção de conhecimento sobre a formação de professores em práticas avaliativas envolvendo inteligência artificial generativa.

Palavras-chave: Inteligência artificial generativa; formação de professores; Química.

Abstract

This work is the result of research that aimed to investigate, from the perspective of undergraduate chemistry students, the contributions of ChatGPT in the development of a question to assess the learning of a chemical content covered in high school. Learning assessment is one of the teaching activities, and since it can be developed with the aid of generative artificial intelligence, it is important that undergraduate students reflect on its limitations. To achieve this objective, we developed an assessment question within the scope of the subject Assessment of the Teaching and Learning Process in Chemistry, which is present in the curriculum of undergraduate chemistry programs at a federal university located in Minas Gerais, aligning the theme of assessment with the use of generative artificial intelligence. From the responses of the five undergraduate chemistry students enrolled in the subject, we constructed the data and interpreted it through content analysis. We found that the students identified that ChatGPT provided answers that need to be reviewed and corrected. This indicates that future teachers will not use the information obtained without verifying its accuracy, which suggests a

¹ Universidade Federal de Juiz de Fora – UFJF. Juiz de Fora/MG, Brasil. E-mail: andreia.afonso@ufjf.br
Orcid: <https://orcid.org/0000-0002-0107-2490>

² Universidade Federal de Juiz de Fora – UFJF. Juiz de Fora/MG, Brasil.
E-mail: guto.carvalho@estudante.ufjf.br - Orcid: <https://orcid.org/0000-0003-3698-1982>

learning process for teaching. Given the results obtained, we emphasize the importance of conducting further research, as there is still a gap in knowledge production regarding teacher training in assessment practices involving generative artificial intelligence.

Keywords - Generative artificial intelligence; teacher training; Chemistry.

Resumen

Este trabajo es el resultado de una investigación que tuvo como objetivo investigar, desde la perspectiva de estudiantes de química de grado, las contribuciones de ChatGPT en el desarrollo de una pregunta para evaluar el aprendizaje de un contenido químico impartido en la escuela secundaria. La evaluación del aprendizaje es una de las actividades docentes y, dado que puede desarrollarse con la ayuda de la inteligencia artificial generativa, es importante que los estudiantes de grado reflexionen sobre sus limitaciones. Para lograr este objetivo, desarrollamos una pregunta de evaluación en el marco de la asignatura "Evaluación del Proceso de Enseñanza y Aprendizaje en Química", presente en el currículo de los programas de química de grado de una universidad federal ubicada en Minas Gerais, alineando el tema de la evaluación con el uso de la inteligencia artificial generativa. A partir de las respuestas de los cinco estudiantes de química de grado matriculados en la asignatura, construimos los datos y los interpretamos mediante análisis de contenido. Observamos que los estudiantes identificaron que ChatGPT proporcionó respuestas que requieren revisión y corrección. Esto indica que los futuros docentes no utilizarán la información obtenida sin verificar su precisión, lo que sugiere un proceso de aprendizaje para la enseñanza. Dados los resultados obtenidos, destacamos la importancia de realizar más investigaciones, ya que aún existe un vacío en la producción de conocimiento respecto a la formación docente en prácticas de evaluación que involucran inteligencia artificial generativa.

Palabras clave - Inteligencia artificial generativa; formación del profesorado; Química.

Introdução

A avaliação é uma ação inerente a diferentes ações presentes no cotidiano. No âmbito educacional, ela é essencial para o acompanhamento da aprendizagem, para a verificação do ensino, para a análise do currículo e para a tomada de atitude diante das dificuldades que são apresentadas pelos estudantes (Luckesi, 2011). Diante dessas finalidades, ressaltamos a importância do estudo dessa temática nos cursos de formação inicial de professores frente aos avanços teóricos e práticos, com a chegada de inteligências artificiais generativas (IAG) às salas de aula e a outros espaços de ensino e de aprendizagem.

Contudo, existem dúvidas sobre a diferença entre uma prática avaliativa e uma prática examinativa. Muitas vezes, os exames podem predominar nos contextos escolares (Luckesi, 2011) por serem uma ação mais frequentemente desenvolvida. Silva e Afonso (2025), inclusive, indicam que alguns licenciandos, participantes de sua pesquisa, apesar de indicarem que haviam estudado sobre avaliação da aprendizagem, ainda eram influenciados pelas práticas examinativas pelas quais passaram na Educação Básica e durante o processo

formativo na licenciatura. Isso pode ser justificado, uma vez que diversos formadores de professores vivenciaram a prática examinativa durante todo o processo de formação, sem momentos de reflexão crítica sobre ela, como apontado por Maldaner (2013).

O autor discorre ainda acerca do que denomina Formação Ambiental, pela qual tais docentes foram formados. Esta influência pode fazer com que os exames se perpetuem, sem que se tenha uma identificação real do processo de aprendizagem dos estudantes. A Formação Ambiental, como supracitado, “cria uma ideia demasiado restrita e muito simples da profissão docente” (Maldaner, 2013, p. 289), a qual, por sua vez, segundo Diniz-Pereira (2014, p. 35), trata-se de uma formação relacionada à Racionalidade Técnica, cuja construção da “prática educacional é baseada na aplicação do conhecimento científico e questões educacionais são tratadas como problemas ‘técnicos’, os quais podem ser resolvidos objetivamente por meio de procedimentos racionais da ciência”, propiciando um viés reducionista da ação professoral, e, eventualmente, acarretando na manutenção da prática examinativa, ao invés de seguir uma formação docente que estimule um olhar crítico-reflexivo que possa levar em conta não somente o conteúdo, como também o contexto, o viés histórico e as particularidades dos estudantes, tornando o ensino mais plural, acessível, promovendo uma aprendizagem participativa, ativa, engajada e equânime.

A reflexão na e sobre a prática, como abordada por Schön (1983) e ampliada por Tardif e Moscoso (2018), durante o processo de formação, são fundamentais para a construção de uma prática docente em constante mudança, sempre buscando aprimoramentos no campo teórico e prático, em consonância com a dinamicidade das salas de aula contemporânea, onde estudantes se encontram em constante adaptação aos desenvolvimentos tecnológicos digitais, como a apropriação de IAG. Entretanto:

No contexto atual, a inteligência artificial emerge como um novo desafio e, ao mesmo tempo, como uma possibilidade para o ensino e a aprendizagem. Ferramentas baseadas em IA, como sistemas adaptativos, plataformas de análise de dados educacionais e assistentes virtuais, ampliam as possibilidades de personalização do ensino, exigindo dos docentes, competências específicas para sua utilização crítica e responsável (Neves et al., 2026, p. 4).

Apesar das dificuldades, a apropriação do uso da IAG é em prol de facilitar a aprendizagem, bem como desenvolver modos de avaliação personalizados, céleres e

eficientes, no intuito de propiciar aos estudantes um feedback de maior qualidade e que possa instigá-los a participarem de modo mais ativo no processo de construção do conhecimento. Além disso, ela pode ser uma ferramenta de geração automática de textos, correção, síntese e recomendação de conteúdos, passando a atuar como coautora nos processos de produção discursiva (Silva et al., 2026, p. 2).

O estudo da avaliação envolve diferentes fatores: planejamento, definição de objetivos, escolha dos instrumentos, estabelecimento de critérios para correção (também denominadas rubricas), estruturação das questões e seleção dos níveis cognitivos da Taxonomia de Bloom Revisada (Ferraz; Belhot, 2010). Sobre a Taxonomia, apontamos que esta é um meio de guiar o percurso cognitivo do estudante, que deve ser progressivo e coerente com sua maturidade, porém, não imprescindível para a elaboração de questões.

Em relação aos instrumentos avaliativos a serem utilizados, ressaltamos o auxílio das tecnologias nesse processo, entre elas a IAG. Rodrigues e Rodrigues (2023, p. 2) afirmam que o entrelaçamento das tecnologias nas Ciências Humanas “requer o entendimento de que as práticas sociais, suas relações com novos meios e mídias é que moldam novas linguagens e novas relações”, abrindo espaço para, também, repensarmos os aspectos e modos de avaliar e construir os instrumentos que lançaremos mão para os devidos fins avaliativos. Neste sentido, as autoras complementam, indicando que “a IA pode otimizar criações, inovações, comparações de dados com maiores tendências” (Rodrigues; Rodrigues, 2023, p. 4), o que pode auxiliar os professores a melhor estruturar o processo de avaliação.

No entanto, as mesmas autoras pontuam que “a preocupação deve dar primazia à qualidade e à veracidade das informações, além da capacidade crítica dentro do contexto em que se aplica o uso de ferramentas, como o *ChatGPT*” (Rodrigues; Rodrigues, 2023, p. 4). Fonseca et al. (2026, p. 3) destacam outra preocupação: “A IA surge como uma solução potencial para automatizar e aprimorar esse processo³, porém, sua implementação ainda enfrenta desafios técnicos, pedagógicos e ético”.

Rudolph, Tan e Tan (2023), em diálogo com outros autores, salientam que, para além de facilitar trabalhos administrativos e burocráticos, o uso de IAG na Educação propicia uma ampliação das ações docentes, a saber: maior atenção dos estudantes e transformação do processo de ensino, bem como, aqui em destaque, o processo de avaliação, de tal modo a

³ Elaboração de questões e avaliações educacionais.

superar “os métodos de avaliação convencionais, os sistemas de avaliação com tecnologia de IA podem dar suporte à integração de feedback contínuo” (p. 351, tradução nossa), o que traria celeridade nas intervenções pedagógicas em prol de melhor atender e orientar os estudantes, bem como realizar tais ações de forma personalizada (Bernius; Krusche; Bruegge, 2022). Além disso, é importante salientar que “o feedback ajuda os alunos a compreender seu progresso na aprendizagem e auxilia no processo de aprendizagem por meio da reflexão” (Bernius; Krusche; Bruegge, 2022, p. 3, tradução nossa).

A partir do exposto, realizamos uma investigação, que foi guiada pela seguinte questão: Como os licenciandos em Química avaliam as contribuições do ChatGPT na elaboração de uma questão para a avaliação da aprendizagem? Para respondê-la, definimos como objetivo da pesquisa: Investigar, na concepção dos licenciandos em Química, as contribuições e limitações do ChatGPT na elaboração de uma questão, para avaliar a aprendizagem de um conteúdo químico, abordado no Ensino Médio.

O ChatGPT é um sistema desenvolvido pela OpenAI, que utiliza a tecnologia de inteligência artificial para fornecer respostas e interagir em conversas com os usuários. Ele é baseado na arquitetura GPT (Generative Pre-trained Transformer), que é um modelo de linguagem neural treinado em grandes quantidades de dados textuais (Rocha; Brandão; Brandão, 2024, p. 5).

Para apresentar os resultados obtidos, estruturamos o artigo nas seguintes seções: Introdução, com a apresentação da temática estudada; Percurso Metodológico, com a descrição das etapas de obtenção e interpretação dos dados; Resultados e Discussão, na qual os resultados postos em evidência e em diálogo com referenciais teóricos pertinentes à temática; Considerações Finais, com uma conclusão sobre a pesquisa realizada e as lacunas ainda existentes e, por fim, as Referências Bibliográficas utilizadas.

Percurso metodológico

Para o desenvolvimento da pesquisa, adotamos a abordagem qualitativa, uma vez que ela “é uma abordagem fundamental na investigação científica, que se baseia na compreensão aprofundada e na interpretação dos fenômenos estudados” (Guerra et al., 2024, p. 3). Além disso:

Os fundamentos da pesquisa qualitativa estão ancorados em princípios teóricos e metodológicos que orientam a coleta e a análise dos dados. Entre os principais fundamentos dessa abordagem estão a busca pela compreensão contextualizada dos fenômenos, a valorização da subjetividade e da diversidade de perspectivas, e a ênfase na flexibilidade e na adaptabilidade do processo de pesquisa (Guerra et al., 2024, p. 3).

Além disso, a pesquisa também pode ser classificada como exploratória

As pesquisas exploratórias vêm sendo utilizadas cada vez mais para investigar os fenômenos complexos da realidade educacional. Esse tipo de investigação busca respostas para questionamentos e dedica-se a identificar e compreender fatos/acidentes da educação que precisam ser explorados. Não se trata de uma simples consulta popular, o propósito é envolver o sujeito que participará desse processo de investigação em um momento de reflexão, análise da realidade e produção de conhecimento (Lösch; Rambo; Ferreira, 2023, p. 3).

A pesquisa aconteceu no âmbito da disciplina Avaliação do Processo de Ensino e Aprendizagem de Química, que faz parte do currículo de três cursos de Licenciatura em Química (integral, noturno e à distância) de uma universidade federal, localizada em Minas Gerais, desde 2020. Nos cursos presenciais, ela é ofertada no sexto período. Já na modalidade EaD, aparece no quinto período, é acontece de forma totalmente assíncrona. Neste artigo, nosso foco foram as aulas presenciais, das quais participaram os licenciandos em Química dos cursos integral e noturno.

As aulas totalizam 30 horas e os licenciandos em Química têm a oportunidade de estudar, refletir e discutir sobre a

[...] avaliação da aprendizagem no âmbito escolar e durante o processo de formação inicial de professores; Apresentação, discussão e elaboração de Instrumentos avaliativos para aulas de Química; Compreensão das avaliações de larga escala e suas respectivas Matrizes de referência; Análise dos resultados obtidos nas avaliações de larga externa (UFJF, 2019, s/p).

Da primeira à quarta aula, levantamos as concepções prévias dos licenciandos em Química a respeito dos diferentes tipos de avaliações, dos sentimentos que surgem diante de lembranças de práticas avaliativas e leituras de referenciais teóricos com posterior discussão. Na quinta aula da disciplina, ofertada no primeiro semestre letivo de 2023, a professora e o

estagiário docente⁴ (autores desse trabalho) propuseram aos cinco licenciandos matriculados, que elaborassem uma questão para avaliar a aprendizagem de Química de uma turma do Ensino Médio. De acordo com Berbel (2001), a oportunidade de refletir sobre a prática avaliativa é uma necessidade formativa dos cursos de licenciatura.

Muitas vezes, a elaboração de questões avaliativas é considerada simples, mas quando os licenciandos foram convidados a redigirem questões originais, ou seja, sem que fossem copiadas de qualquer fonte, as dificuldades surgiram. Uma delas se referiu ao domínio do conteúdo químico a ser avaliado e a outra foi a redação do enunciado de forma que fosse compreensível para que o objetivo determinado para a questão fosse alcançado, ou seja, para que o estudante conseguisse mostrar o conhecimento construído até o momento e o docente pudesse identificar as possíveis dificuldades de aprendizagem ainda presentes.

Na sexta aula, os cinco licenciandos foram convidados a apresentarem suas questões, de forma que refletissem sobre a potencialidade e a fragilidade das suas criações, pois, algumas vezes, a estrutura e o comando da questão parecem estar claros para quem os redigi, mas não o é para aqueles que buscam resolvê-la. Portanto, esse encontro também foi o momento para a reescrita das questões a partir das sugestões dos colegas, da professora e do estagiário.

Uma vez sanadas as possíveis dificuldades encontradas, na oitava aula, a professora regente aplicou um Teste de Verificação do Conhecimento (TVC)⁵ com três questões, das quais uma delas se tornou o objeto de nosso estudo. A referida questão tinha a seguinte redação: “Elabore o mesmo item⁶ no ChatGPT indicando as diferenças entre aquele que você elaborou e o produzido pela “inteligência artificial. Faça uma análise crítica de todos os aspectos que identificou no item produzido pelo ChatGPT”.

Embora seja capaz de gerar respostas relevantes, em certas situações pode produzir informações incorretas, irrelevantes ou não compreender completamente o contexto da conversa. Além disso, o sistema pode ser sensível a vieses presentes nos dados de treinamento, o que pode influenciar suas respostas (Rocha; Brandão; Brandão, 2024, p. 5).

⁴ Doutorando da área de Educação Química, que acompanhou a disciplina para cumprir o crédito de Estágio Docência.

⁵ São avaliações aplicadas ao longo das disciplinas e podem ser realizadas por meio de diferentes instrumentos.

⁶ Esse item se refere ao item reescrito na sexta aula, a partir das considerações dos colegas de turma da professora e do estagiário, como descrito no parágrafo anterior.

Por isso, essa questão tinha como finalidade estabelecer um paralelo entre o item original (criado pelos licenciandos) e o item criado a partir da IAG, promovendo a reflexão crítica sobre as limitações e contribuições da tecnologia. Portanto, essa questão do TVC estava relacionada à inteligência artificial generativa (IAG), temática discutida em algumas aulas da disciplina, uma vez que os futuros professores podem se apropriar de questões retiradas da internet para serem incorporadas em suas avaliações, por vários motivos, sem a devida análise crítica (que era a proposta da questão), já que:

O ChatGPT é resultado de anos de pesquisa em inteligência artificial e aprendizado de máquina, e é capaz de responder a uma ampla variedade de perguntas em linguagem natural, desde questões simples até perguntas mais complexas e técnicas. Ele é treinado em uma grande quantidade de dados de texto em vários idiomas, o que o torna capaz de entender e produzir texto em diferentes idiomas (Monteiro, 2023, p. 2901).

O TVC foi disponibilizado no *Google Classroom* e os cinco licenciandos tiveram uma semana para postarem suas respostas no formato *word* ou *pdf*. Realizamos *download* das respostas, já que elas se constituíram nosso *corpus* de análise e iniciamos a leitura minuciosa, destacando trechos considerados essenciais para a investigação. Esses trechos, segundo a Análise de Conteúdo (Bardin, 2016), são denominadas unidades de registro. Elas foram codificadas com a letra L (de licenciando), seguida de uma numeração em ordem crescente, de acordo com o posicionamento no texto. Por exemplo: Unidade L1.2

L = Licenciando em Química.

1 = Primeiro licenciando da lista, que foi organizada em ordem alfabética.

2 = Segundo trecho destacado da resposta de L1.

As unidades de registro foram organizadas em categorias que permitiram interpretar os resultados, como será mostrado na próxima seção. As três categorias – Contribuições, Limitações e Aprendizagens para a Docência - foram criadas *a posteriori* a partir da leitura minuciosa do material analisado, do objetivo determinado e do estudo dos referenciais teóricos referentes à temática da pesquisa.

Vale destacar que as identidades dos licenciandos em Química foram mantidas sob sigilo e, por isso, seus nomes foram substituídos por outros fictícios.

Resultados e discussão

As respostas dos cinco licenciandos não apresentaram uma análise aprofundada, como se esperava. Esse fato pode estar associado a dificuldade na escrita reflexiva, diferente daquela que estão habituados, quando redigem os relatórios dos laboratórios (Araújo, 2024), que é mais objetiva, sem a expressão de suas reflexões.

Entretanto, identificamos alguns aspectos, discutidos nas aulas, que também foram apontados e criticados durante as apresentações. Flávia, por exemplo, fez sua análise da seguinte forma:

Comparado com o item⁷ criado por mim, ele [item criado a partir do ChatGPT] não possui texto base, somente uma frase para explicar do que se trata a questão. Assim como o meu item, teve a ausência de um comando, ou senão, um comando em forma de pergunta.

A partir da resposta de Flávia, podemos perceber que a sua análise se deu em função da estrutura da questão: presença (ou não) de texto base, comando e enunciado. A estrutura foi apresentada em aulas anteriores, o que pode ter influenciado a sua percepção sobre a questão criada pelo ChatGPT. Contudo, nosso objetivo principal era que Flávia refletisse sobre o conteúdo, a redação do enunciado e das possíveis alternativas, caso fosse criada uma questão objetiva.

A redação da questão é importante, pois segundo Luckesi (2018), os instrumentos avaliativos precisam viabilizar uma adequada descrição da realidade da aprendizagem. Caso contrário, o mau desempenho pode estar associado a dificuldade na interpretação e compreensão da questão e não, necessariamente, ao conhecimento que está sendo avaliado.

Além disso, a licencianda parece ter tido uma preocupação com a contextualização: “O item também não apresenta nenhuma problematização que possa ser solucionada pela questão, somente o conteúdo pelo conteúdo, sem uma associação com algum conteúdo do cotidiano do aluno”. Sara também ficou atenta a contextualização e solicitou ao ChatGPT uma questão, nessa perspectiva, sobre transformações químicas e físicas por considerar que exemplos cotidianos aparecem com mais frequência nessa temática.

⁷ Neste artigo, estamos utilizando item e questão como sinônimos, embora reconheçamos que existem diferenças nas duas terminologias.

Entretanto, a licencianda relatou ter encontrado erros conceituais nas respostas, assim como Murilo, e relata: “Ainda não satisfeita com a contextualização, mudei a ordem das palavras: Faça uma questão contextualizada de prova de Química em que seja necessário saber diferenciar transformações químicas e físicas”. A inquietação de Sara no leva a acreditar que houve uma primeira análise crítica da questão criada pela IAG, uma vez que ela prosseguiu, indicando um outro *input*.

Mas os erros conceituais persistiram nas alternativas (o ChatGPT criou uma questão objetiva com quatro alternativas). Diante deles, Sara questionou a IAG: “Tem certeza dessa resposta?”. Esse questionamento nos permite afirmar que a licencianda foi crítica, mas uma vez, em relação a utilização da tecnologia. Segundo ela:

Isso mostrou o que temos conversado em sala de aula, sobre questionar as informações que uma “inteligência artificial” nos dá. É possível utilizá-la para produzir itens, mas, se queremos algo contextualizado (seja dentro da realidade de uma região ou geral), é melhor que seja um trabalho nosso. Além disso, utilizar a nossa parte inteligente para ver se o que está escrito está correto, pois pode não estar.

Fato que vai ao encontro da conclusão de Rodrigues e Rodrigues (2023, p. 10), quando as autoras apontam de que o uso de IAG não assevera a obtenção dos objetivos do processo por meio de máquina que esteja “supostamente protegida contra erros humanos, fator observado no uso da ferramenta ChatGPT, cujas buscas e contribuições não estão isentas de erros”, o que exige um olhar atento e crítico, em consonância com o aspecto reflexivo, e tendo em conta o âmbito contextual e social.

Desse modo, por terem apresentado o comando dado ao ChatGPT e terem refletido sobre ele, Natália, Sara, Gisele e Murilo foram mais detalhistas, quando comparamos suas respostas com as de Flávia, pois os últimos quatro apresentaram de que forma enviaram o comando para o ChatGPT:

Descreva como as células eletroquímicas são utilizadas em baterias recarregáveis contemporâneas, como aquelas encontradas em telefones celulares e laptops. Analise o processo de carregamento e descarregamento das células eletroquímicas nestas baterias, destacando os componentes e circuitos envolvidos no controle de carga e na preservação da vida útil da bateria (Natália).

Minha primeira tentativa foi associar a habilidade reescrita: “Faça uma questão de prova em que seja necessário diferenciar transformações e

Revista *Devir Educação*, Lavras, vol.10, n.1, e-1173, 2026.

conservações em sistemas que envolvam quantidade de matéria, de energia e de movimento” (Sara).

Função horária da posição em um experimento de queda livre (Gisele).

Elabore uma questão de Química para o Ensino Médio sobre as baterias de chumbo/óxido de chumbo (chumbo/ácido), envolvendo os conteúdos de células eletroquímicas, equilíbrio químico e reações reversíveis. Estes conteúdos devem ser articulados de modo que os alunos devam ser submetidos ao nível cognitivo de Análise, conforme a Taxonomia de Bloom. Para solucionar a questão, os alunos devem demonstrar a habilidade de analisar os processos de transformação da matéria e condução de energia envolvidos no funcionamento de pilhas e baterias, por meio de previsões qualitativas ou quantitativas, com ou sem o uso de dispositivos e aplicativos digitais, para compreensão das tecnologias contemporâneas (Murilo).

Cabe destacar que Murilo utilizou vários comandos (*input*), indicando que gostaria que o ChatGPT apresentasse uma questão completa, que abarcasse todos os objetivos e conteúdos propostos por ele. Essa ação vai ao encontro do trabalho de Silva et al. (2026, p. 7), para os quais, “a escrita com apoio da IA não pode ser reduzida a um atalho para a obtenção de respostas prontas, sob pena de esvaziar seu potencial formativo”.

O último comando estava relacionado a verificação das fontes bibliográficas utilizadas pela IAG, uma vez que, dessa forma, o licenciando poderia se certificar se o que foi indicado poderia ser utilizado ou não em uma avaliação. A “utilização de IA pode garantir uma maior diversidade e equidade nas avaliações, uma vez que é possível evitar preconceitos e assegurar que todas as questões sejam justas para todos os grupos de alunos” (Fonseca et al., 2026, p. 4).

Assim como Murilo, Sara também teve o cuidado de verificar, solicitando ao ChatGPT, a resposta da questão que formulou. Segundo Sara, isso foi preciso para que pudesse fazer a análise de forma mais adequada. Podemos inferir que esses momentos também proporcionaram o estudo dos conceitos abordados na avaliação que estavam elaborando.

Já Natália considerou a questão apresentada pelo ChatGPT completa, por isso, não viu necessidade de verificar a resposta. Essa percepção se dá pelo fato da questão envolver não só a mobilização do conhecimento do estudante, relacionado a funcionalidade das baterias recarregáveis, como também dos processos envolvidos nas pilhas e baterias, sendo eles espontâneos ou não,

[...] o que me leva a questionar como o aluno do Ensino Médio, por exemplo, conseguiria dentro de tudo que estudamos na eletroquímica, associar somente energia elétrica e energia química na bateria, sem passar por todos os fatores como espontaneidade, os polos negativos e positivos ligados para gerar energia elétrica, e como o professor vai poder analisar o processo de aprendizagem se nem mesmo o site [ChatGPT] responde com clareza o que é perguntado. Esse tipo de questão não me deixa avaliar todo o entendimento do aluno sobre o tema abordado, pois vai além do que é trabalhado em sala de aula.

Natália levanta um aspecto interessante e importante quando se trata de avaliação da aprendizagem: a capacidade da questão elaborada de identificar o conhecimento construído pelo estudante que está sendo avaliado, o que vai ao encontro do confronto de Sara com o ChatGPT. Para Natália, a questão criada pelo ChatGPT pode falhar na identificação das aprendizagens dos estudantes, pois até mesmo a IAG não conseguiu respondê-la da forma que se esperava.

Para a IA, o que deve ser avaliado são os educandos, a fim de verificar suas compreensões dos conteúdos ensinados. Mas, como já foi supracitado, a avaliação deve ser contínua, de modo a avaliar todo o processo de ensino e aprendizagem, e não os sujeitos aprendizes (Rocha; Brandão; Brandão, 2024, p. 4).

A verificação feita pelas licenciandos foi importante para revalidar as afirmações de Luckesi (2011) sobre a importância de ter objetivos bem definidos durante o planejamento das avaliações da aprendizagem. Além disso, de acordo com Monteiro (2023, p. 2904),

[...] é importante destacar que o uso do ChatGPT deve ser planejado cuidadosamente para garantir que questões éticas, como privacidade, segurança e transparência, sejam levadas em consideração. É necessário que os educadores estejam atentos à forma como utilizam a tecnologia e que os alunos sejam orientados a usá-la de maneira crítica e responsável.

A mesma preocupação de Natália foi identificada na resposta de Sara, que afirmou não ter ficado satisfeita com a questão criada pelo ChatGPT, que gerou uma sentença muito extensa, envolvendo conteúdos além daqueles que ela pensou avaliar os estudantes do Ensino Médio. Natália e Sara parecem concordar que a aprendizagem se dá por meio de uma associação de conceitos previamente estudados, e por terem escolhido uma temática que envolve vários deles, a questão deve ser detalhada e elaborada de forma clara e cuidadosa

(Luckesi, 2011), de forma a permitir que o estudante consiga expressar o conhecimento relacionado a eletroquímica.

Para resolver isso, Sara solicitou à IAG: “Faça uma questão de prova de Química em que seja necessário saber diferenciar transformações químicas e físicas”. Dessa vez, a licencianda considerou que a questão está mais próxima dos objetivos propostos por ela:

- O aluno deve ser capaz de identificar corretamente se cada transformação é química ou física e justificar sua resposta.
- O aluno deve ser capaz de explicar a diferença entre transformações químicas e físicas e dar um exemplo de cada uma.
- O aluno deve ser capaz de interpretar a tabela de variações de entalpia (ΔH) e classificar cada transformação como endotérmica ou exotérmica (Sara).

Para Luckesi (2018), é preciso planejamento e definição dos objetivos da prática avaliativa, etapas do “projeto de ação investigativa” (p. 49). Caso contrário, os resultados não retratarão a realidade que se quer investigar e as dificuldades na aprendizagem poderão persistir.

Uma outra diferença relatada por Gisele, ao comparar a sua questão com a criada pelo ChatGPT, está relacionada a representação. Na questão criada por ela, há uma imagem por meio da qual o estudante pode acompanhar a situação dada no enunciado. Já na questão da IAG, o estudante é convidado a responder à questão de forma experimental e comparar os dados obtidos antes da realização da prática. Assim, Gisele avalia a questão do ChatGPT: “De uma forma geral, é um item bastante rico para ser trabalhado, pois ele atende muito bem as habilidades e indo além delas um pouco, mas ao mesmo tempo, não saindo delas”. As habilidades foram extraídas da Base Nacional Comum Curricular e reescritas em uma atividade anterior.

A representação também foi um aspecto citado por Murilo:

Eu considero que o item elaborado por mim está melhor fundamentado, pois a linguagem que eu utilizo: texto, imagem, e a maneira como são apresentadas, exigem sobretudo que o aluno seja capaz de analisar a situação-problema, mobilizando o conhecimento químico. Com relação a habilidade reescrita, o que proponho está em melhor conformidade do que o item apresentado pela inteligência artificial.

Murilo foi um pouco mais além nas suas respostas, quando direciona sua análise para o nível cognitivo, de acordo com a Taxonomia de Bloom Revisada (Ferraz; Belhot, 2010). De acordo com ele:

Neste sentido, o êxito na resolução do problema é fundamentado em um nível cognitivo inferior altamente conteudista (LEMBRAR), de modo que o aluno precisaria conhecer todas as reações químicas envolvidas na bateria de chumbo/ácido presente no motor de um carro. Ratifico que para o domínio de um nível cognitivo ANALISAR, os outros níveis inferiores também devem ser dominados de modo satisfatório pelo aluno, todavia, o item deve ser direcionado para o nível cognitivo condizente com a habilidade a ser trabalhada, o que não foi feito pelo ChatGPT.

A Taxonomia de Bloom Revisada (Ferraz; Belhot, 2010) foi um dos pontos discutido nas aulas da disciplina Avaliação do Processo de Ensino e Aprendizagem de Química. Pela resposta, verificamos que Murilo teve uma preocupação com o desenvolvimento cognitivo, de forma que as questões pudessem indicar qual nível precisaria de mais atenção por parte do professor, de forma que a progressão aconteça de forma progressiva e coerente com a habilidade trabalhada nas aulas.

Nesse sentido, os cinco licenciandos em Química demonstraram ter mobilizado diferentes temáticas, abordadas durante as aulas, que os auxiliaram na análise crítica das questões criadas por eles. Esse pode ter sido o início de uma aprendizagem a ser levada para a futura prática avaliativa, tão importante quanto as outras atividades próprias da docência.

Além disso, esses futuros professores demonstraram uma percepção crítica-reflexiva do uso de recursos desenvolvidos pela IAG, como o ChatGPT, a fim de melhorar o processo de avaliação, tendo mecanismos mais eficientes no intuito de averiguar as potencialidades e fragilidades do(s) método(s) de ensino utilizado(s) para lecionar determinado(s) conteúdo(s).

Considerações finais

A prática avaliativa é uma das atividades docentes que envolve o acompanhamento da aprendizagem, a verificação do ensino, a análise do currículo e o planejamento de ações que visam superar as dificuldades apresentadas pelos estudantes. Por isso, a formação, voltada para essa temática, é essencial, especialmente, quando promove a reflexão, o estudo e o desenvolvimento de instrumentos que cumpram os objetivos de uma avaliação de qualidade.

Revista *Devir Educação*, Lavras, vol.10, n.1, e-1173, 2026.

Não podemos esquecer que as tecnologias, principalmente a inteligência artificial generativa, estão disponíveis e podem auxiliar em diferentes tarefas. Os licenciandos em Química, participantes da pesquisa, a utilizam com frequência em suas atividades e, por isso, é necessária uma análise crítica sobre os resultados que se obtêm por meio dela. Quando essa análise está alinhada à avaliação, as contribuições podem ser potencializadas.

No estudo realizado, verificamos que os licenciandos identificaram aspectos, que precisam ser revistos e corrigidos, a partir das respostas obtidas por meio do ChatGPT. Isso indica que os futuros professores não utilizarão as informações da IAG sem se certificar que estão corretas, o que indica que a criticidade esteve presente, assim como a preocupação com a criação de um instrumento avaliativo sem falhas.

O uso de IAG deve ser articulado por um viés crítico-reflexivo, sobretudo quando se deseja ter um processo de ensino articulado com estas tecnologias, a fim de se construir processos avaliativos que auxiliem os professores a identificarem os eventuais erros dos estudantes e, dessa forma, poderem melhor orientar estes estudantes a estudarem, obtendo níveis cognitivos mais avançados e superarem os obstáculos, desenvolvendo modos de construção de conhecimento mais eficazes.

Por fim, embora existam muitas pesquisas sobre IAG, outras relacionadas à avaliação precisam ser produzidas. Ainda há uma lacuna na produção de conhecimento sobre a formação de professores em práticas avaliativas envolvendo IAG. Que este artigo seja um propulsor e incentivador para a intensificação de outros estudos, para que as discussões sejam ampliadas e aprofundadas.

Referências

ARAÚJO, Larissa Dias da Silveira. **O Diário de Bordo na formação inicial de professores de Química:** concepções, possibilidades e desafios. 2024. 116 p. Dissertação (Mestrado em Química) – Universidade Federal de Juiz de Fora/UFJF, Juiz de Fora, 2024.

BARDIN, Laurence. **Análise de conteúdo.** Edições 70, 2016.

BERBEL, Neusi Aparecida Navas. Dimensão Pedagógica. In: BERBEL, Neusi Aparecida Navas; COSTA Waldecir Souza da; GOMES, Icléia Rodrigues de Lima; OLIVEIRA, Cláudia Chueire de; VASCONCELLOS, Maura Maria Morita. **Avaliação da Aprendizagem no Ensino Superior:** Um retrato em cinco dimensões. Londrina: UEL, 2001.

BERNIUS, Jan Philip; KRUSCHE, Stephan; BRUEGGE, Bernd. Machine Learning based feedback on textual student answers in large courses. **Computers and Educations: Artificial Intelligence**, v. 3, p. 1-16, 2022.

DINIZ-PEREIRA, Júlio Emílio. Da racionalidade técnica à racionalidade crítica: formação docente e transformação social. **Perspectivas Em Diálogo Revista de Educação e Sociedade**, v.1, n. 1, p. 34-42, jan-jun.2014.

FERRAZ, Ana Paula do Carmo Marcheti; BELHOT, Renato Vairo. Taxonomia de Bloom: revisão teórica e apresentação das adequações do instrumento para definição de objetivos instrucionais. **Gestão & Produção**, São Carlos, v. 17, n. 2, p. 421-431, 2010.

FONSECA, Andreza Bastos Bartz Nogueira da; ARAÚJO, Andréia Freitas de; CHAVES, Alessandra Encinas Brandão; LONGO, Conceição dos Santos; FERREIRA, Herberth Gomes; SANTOS, Luciano Gabriel dos. Inteligência artificial na geração de questões e avaliações educacionais. **Revista Ibero-Americana de Humanidades, Ciências e Educação**, São Paulo, v. 12, n.1, p. 1-7, 2026.

GUERRA, Avaetê de Lunetta e Rodrigues; STROPARO, Telma Regina; COSTA, Michel da; CASTRO JÚNIOR, Francisco Pires de; LACERDA JÚNIOR, Orivaldo da Silva; BRASIL, Melca Moura; CAMBA, Mariangela. Pesquisa qualitativa e seus fundamentos na investigação científica. **Revista de Gestão e Secretariado**, v. 15, n. 7, p.1-15, 2024.

LÖSCH, Silmara; RAMBO; Carlos Alberto; FERREIRA, Jacques de Lima. A pesquisa exploratória na abordagem qualitativa em Educação. **Revista Ibero-Americana de Estudos em Educação**, Araraquara, v. 18, n. 00, e023141, p. 1-18, 2023.

LUCKESI, Cipriano Carlos. **Avaliação da aprendizagem**: componente do ato pedagógico. 1 ed. São Paulo: Cortez, 2011.

LUCKESI, Cipriano Carlos. **Avaliação em Educação**. São Paulo: Cortez, 2018.

MALDANER, Otavio Aloisio. A formação inicial e continuada de professores de química: professor/pesquisador. **Editora Unijuí**, 4ª ed, 422 p., 2013.

MONTEIRO, Jean Carlos da Silva. Assistente CHATGPT na educação: possibilidades e desafios. **Revista Ibero-Americana de Humanidades**, 9(6), p.2899–2906, 2023.

ROCHA, Telma Brito; BRANDÃO, Cleyton Williams Golveia da Silva; BRANDÃO, Williams Panfile Santos. Avaliação da aprendizagem digital segundo o ChatGPT-3.5 e a autoria híbrida humano - IA. **AtoZ: novas práticas em informação e conhecimento**, 13, p. 1-11, 2024.

NEVES, Lucimari Sara das; MENEZES, Nilton Cezar Rodrigues; SILVA, Marilene Soares da; DANTAS, Adriana Doroteu; MEDEIROS, Mailana Dantas; HORTA, Diego Aquino; CUNHA, Francisco Bentes; BANDEIRA, Herick Simon Nunes; SANTOS, Jorge Alberto da Silva; SOUSA, Benedito Salazar; CRUZ, Érico Borgonove; SILVA, Rodrigo Fernandes da; REBOUÇAS, Fábio Augusto Costa Ferreira; RAMOS, Jorge Henrique. Formação docente,

letramento digital e IA: perspectivas para o ensino e aprendizagem. **Revista DCS**, v. 23, n. 86, p. 01-14, 2026.

RODRIGUES, Olira Saraiva; RODRIGUES, Karoline Santos. A inteligência artificial na educação: os desafios do ChatGPT. **Texto Livre**, v. 16, p. 1-12, 2023

RUDOLPH, Jürgen; TAN, Samson; TAN, Shannon. ChatGPT: bullshit spewer or the end of traditional assessments in higher education? **Journal of Applied Learning & Teaching**, vol. 06, nº 1, p. 342-363, 2023.

SCHÖN, Donald. **The reflective practitioner**. New York: Basic Books, 1983.

SILVA, Isabela Vieira; AFONSO, Andréia Francisco. Avaliação da aprendizagem: um estudo sobre os conhecimentos construídos pelos licenciandos em Química ao longo do processo formativo. **Revista Ponto de Vista**, v. 14, n. 1, p.1-18, 2025.

SILVA, Marcus Vinícius da; CARVALHO, Zulene dos Santos; SANTANA, Jacineide Virgínia Borges Oliveira da Silva; LUZ, Raimundo Gomes; MARSAIOLI, Rosiane Almeida Minet; CAMPOS, Simone Batista; XAVIER, Lidianne da Silva; OLIVEIRA, Relinaldo Pinho de; MACHADO, Leandro Soares; OLIVEIRA, Andreia Vanessa de; COSTA, Ramon Santos; SANTOS, Maria do Livramento da Silva; PINTO, Elinor Rangel; ARAÚJO, Alessandra Cristina de. Inteligência artificial na sala de aula: práticas pedagógicas, linguagem e formação crítica. **Revista Ibero-Americana de Humanidades, Ciências e Educação**, São Paulo, v. 12, n.1, jan., p. 1-13, 2026.

TARDIF, Maurice; MOSCOSO, Javier Nunez. A noção de “profissional reflexivo” na educação: atualidade, usos e limites. **Cadernos De Pesquisa**, 48 (168), 388–411, 2018.

UNIVERSIDADE FEDERAL DE JUIZ DE FORA. Departamento de Química. Plano de Ensino. Disponível em: <https://www2.ufjf.br/quimica/disciplinas/plano-departamental/> Acesso em 09 ago. 2025.

Recebido: novembro/2025
Revisões requeridas: fevereiro/2026
Aprovado: fevereiro/2026