



UNIVERSIDADE
ESTADUAL de LONDRINA

SANDRO LUCAS REIS COSTA

**AVALIAÇÃO NA ÁREA DE ENSINO DE QUÍMICA:
CARACTERÍSTICAS, TENDÊNCIAS E LACUNAS NAS
PESQUISAS**

SANDRO LUCAS REIS COSTA

**AVALIAÇÃO NA ÁREA DE ENSINO DE QUÍMICA:
CARACTERÍSTICAS, TENDÊNCIAS E LACUNAS NAS
PESQUISAS**

Tese apresentada ao Programa de Pós-Graduação em Ensino de Ciências e Educação Matemática, da Universidade Estadual de Londrina, como requisito parcial à obtenção do título de doutor em Ensino de Ciências e Educação Matemática.

Orientadora: Profa. Dra. Fabiele Cristiane
Dias Broietti

Londrina
2025

Ficha de identificação da obra elaborada pelo autor, através do Programa de Geração Automática do Sistema de Bibliotecas da UEL

Costa, Sandro Lucas Reis.

Avaliação na Área de Ensino de Química : Características, Tendências e Lacunas nas Pesquisas / Sandro Lucas Reis Costa. - Londrina, 2025.
212 f.

Orientador: Fabiele Cristiane Dias Broietti .

Tese (Doutorado em Ensino de Ciências e Educação Matemática) - Universidade Estadual de Londrina, Centro de Ciências Exatas, Programa de Pós-Graduação em Ensino de Ciências e Educação Matemática, 2025.
Inclui bibliografia.

1. Avaliação - Tese. 2. Revisão Bibliográfica Sistemática - Tese. 3. Ensino de Química - Tese. 4. Pesquisadores - Tese. I. Broietti , Fabiele Cristiane Dias . II. Universidade Estadual de Londrina. Centro de Ciências Exatas. Programa de Pós-Graduação em Ensino de Ciências e Educação Matemática. III. Título.

CDU 37

SANDRO LUCAS REIS COSTA

**AVALIAÇÃO NA ÁREA DE ENSINO DE QUÍMICA:
CARACTERÍSTICAS, TENDÊNCIAS E LACUNAS NAS
PESQUISAS**

Tese de doutorado apresentada ao Programa de Pós-graduação em Ensino de Ciências e Educação Matemática, da Universidade Estadual de Londrina, para a obtenção do título de doutor em Ensino de Ciências e Educação Matemática.

BANCA EXAMINADORA

Orientadora: Profa. Dra. Fabiele Cristiane Dias
Broietti
Universidade Estadual de Londrina - UEL

Prof. Dr. João Paulo Camargo de Lima
Universidade Tecnológica Federal do Paraná –
UTFPR

Profa. Dra. Marinez Meneghello Passos
Universidade Estadual de Londrina - UEL

Profa. Dra. Rosangela Ines Matos Uhmman
Universidade Federal da Fronteira Sul - UFFS

Profa. Dra. Simone Luccas
Universidade Estadual do Norte do Paraná -
UENP

Londrina, 03 de Abril de 2025.

AGRADECIMENTOS

Agradeço à minha orientadora, a professora Fabiele Cristiane Dias Broietti, por enxergar em mim, ainda na graduação, potencial como pesquisador e acadêmico. Por ser sempre uma excelente colaboradora e mentora na elaboração das ideias da tese e na execução das etapas necessárias para torná-las realidade. Ao tempo, energia e esforço que a professora Fabiele investiu em mim como profissional e ser humano ao longo dos últimos 6 anos. Sua receptividade, entusiasmo e paciência impactaram positivamente a minha trajetória e tornaram a experiência do doutorado um marco significativo na minha história de vida.

À professora Marinez Meneghello Passos, por ter me inspirado a ser produtivo e prolífico e sempre buscar novos níveis de excelência e dedicação. Também por seus comentários pontuais e formativos em minhas apresentações no grupo EDUCIM¹.

Ao professor Sergio de Mello Arruda por ter me ensinado a importância de escutar e de sempre estar aberto a novas ideias. Também por mostrar como nosso trabalho como pesquisador é colaborativo e como as relações no trabalho acadêmico podem ser bondosas e gentis.

Aos demais membros da Banca, o professor João Paulo Camargo de Lima, a professora Rosângela Ines Matos Uhmman e a professora Simone Luccas pela sua disposição em participar desta banca e contribuir com seus comentários.

Por fim, aos meus pais e à criação que investiram em mim, os quais inculcaram em mim valores essenciais para a realização desse trabalho.

¹ O grupo EDUCIM – Educação em Ciências e Matemática – foi fundado em 2002 e está associado ao Programa de Pós-Graduação em Ensino de Ciências e Educação Matemática da Universidade Estadual de Londrina (PECEM/UEL). Seu principal objetivo é investigar questões ligadas à formação de professores, bem como ao ensino e à aprendizagem de Ciências e Matemática, tanto no contexto da educação formal quanto à educação informal. A coordenação do grupo é realizada por três orientadores do PECEM: Sergio de Mello Arruda, Marinez Meneghello Passos e Fabiele Cristiane Dias Broietti. Para mais informações: <http://educim.com.br/>

COSTA, Sandro Lucas Reis. **Avaliação na Área de Ensino de Química:** Características, Tendências e Lacunas nas Pesquisas. 2025. 212 f. Tese (Doutorado em Ensino de Ciências e Educação Matemática) – Universidade Estadual de Londrina, Londrina, 2024

RESUMO

Esta pesquisa apresenta resultados de uma Revisão Bibliográfica Sistemática de artigos sobre Avaliação na área de Ensino de Química, bem como resultados de entrevistas com os principais autores internacionais que investigam esta temática abordando as características, as tendências e as lacunas das pesquisas nesse tema. Foram analisados 205 artigos publicados em periódicos internacionais da área de Ensino de Química ao longo de vinte anos (2002-2021), disponíveis em duas bases de dados: ERIC e ERIH Plus. Entrevistou-se quatro autores proeminentes no tema e buscou-se responder às seguintes questões de pesquisa: I) Quais são as principais características, tendências e lacunas dos artigos internacionais sobre Avaliação na área de Ensino de Química? II) O que tem sido avaliado nos artigos internacionais sobre Avaliação na área de Ensino de Química? III) Quais são as perspectivas dos principais autores internacionais em Avaliação na área de Ensino de Química sobre as características, tendências e as lacunas das pesquisas? Para isso, foi realizada uma investigação qualitativa com os procedimentos analíticos orientados pela Análise de Conteúdo de Bardin (2011) e os procedimentos metodológicos conforme o guia para uma revisão bibliográfica sistemática de Okoli (2015). Como resultados, identificou-se um aumento significativo no número de publicações, especialmente entre 2011 e 2021, quando a produção saltou de 18% para 82% do total, indicando uma crescente valorização do tema. Esse crescimento pode ser atribuído à introdução do conceito de avaliação formativa, o aumento de profissionais pesquisando na área de Ensino e uma necessidade crescente de instrumentos avaliativos mais precisos e confiáveis. Identificou-se que os periódicos *Chemistry Education Research and Practice* (66%) e *Journal of Chemical Education* (16%) foram cruciais para a disseminação das pesquisas e a promoção de discussões sobre práticas avaliativas. Os Estados Unidos, mais especificamente o estado da Flórida, destacou-se como um polo de pesquisa, em especial em instituições como a *University of South Florida* e a *Miami University*, onde pesquisadores influentes têm contribuído significativamente para a literatura. Quanto ao nível de ensino avaliado, a maioria dos artigos se concentrou no contexto do ensino superior (50%), enquanto o contexto da Educação Básica representou apenas 8% das publicações. Essa discrepância aponta para uma lacuna na pesquisa que requer maior atenção. A análise do que foi avaliado gerou dez categorias: Avalia habilidades (C01); Avalia estratégias de avaliação, provas, exames e itens (C02); Avalia propostas de ensino (C03); Avalia um instrumento, ferramenta ou tecnologia (C04); Avalia compreensões, percepções e interesse (C05); Avalia conhecimentos (C06); Avalia o currículo (C07); Avalia atitudes (C08); Avalia outras pesquisas (C09); e Discute avaliação em contextos diferentes (C10). Dentre as categorias, a avaliação de propostas de ensino foi a mais proeminente (19%), seguida pela avaliação de instrumentos e tecnologias (17,5%). As perspectivas dos principais

autores internacionais apontam para a necessidade de uma abordagem mais plural nas avaliações, enfatizando o potencial da avaliação na coleta de dados e na realização de inferências sobre a aprendizagem. Entre as lacunas identificadas estão a falta de estudos de longo prazo e a necessidade de compreender melhor as percepções dos alunos e o impacto das práticas de avaliação. Os resultados desta pesquisa revelam uma transformação significativa nas publicações sobre Avaliação na área de Ensino de Química ao longo do recorte temporal e destacam a necessidade de futuras investigações que abordem as lacunas existentes, especialmente em relação à avaliação das práticas docentes e o contexto da Educação Básica. Em futuras investigações, as características, tendências e lacunas apresentadas nesse estudo podem servir como um ponto de partida para promover um entendimento mais abrangente e fundamentado da temática avaliação, em diferentes contextos educacionais.

Palavras-chave: Avaliação. Revisão Bibliográfica Sistemática. Ensino de Química. Pesquisadores. Contexto Internacional.

COSTA, Sandro Lucas Reis. **Assessment in the Field of Chemistry Education: Characteristics, Trends and Gaps in Research**. 2025. 212 p. Thesis (Doctorate in Science Teaching and Mathematics Education) - Londrina State University, Londrina, 2024.

ABSTRACT

This research presents the results of a Systematic Literature Review of articles on Assessment in the field of Chemistry Education, as well as the results of interviews with the main international authors who investigate this topic, addressing the characteristics, trends and gaps in research. A total of 205 articles published in international journals in the field of Chemistry Education over twenty years (2002-2021) were analyzed, available in two databases: ERIC and ERIH Plus. Four prominent authors on the topic were interviewed and sought to answer the following research questions: I) What are the main characteristics, trends and gaps in international articles on Assessment in the field of Chemistry Education? II) What has been evaluated in international articles on Assessment in the field of Chemistry Education? III) What are the perspectives of the main international authors in Assessment in the field of Chemistry Education on the characteristics, trends and gaps in the research? To this end, a qualitative investigation was carried out using analytical procedures guided by Bardin's Content Analysis (2011) and methodological procedures according to Okoli's guide for a systematic literature review (2015). As a result, a significant increase in the number of publications was identified, especially between 2011 and 2021, when production jumped from 18% to 82% of the total, indicating a growing appreciation of the topic. This growth can be attributed to the introduction of the concept of formative assessment, the increase in professionals researching in the field of Education, and a growing need for more accurate and reliable assessment instruments. It was identified that the journals Chemistry Education Research and Practice (66%) and the Journal of Chemical Education (16%) were crucial for the dissemination of research and the promotion of discussions on assessment practices. The United States, more specifically the state of Florida, stood out as a research hub, especially in institutions such as the University of South Florida and Miami University, where influential researchers have contributed significantly to the literature. Regarding the level of education assessed, most articles focused on the context of higher education (50%), while the context of Basic Education represented only 8% of the publications. This discrepancy points to a gap in the research that requires greater attention. The analysis of what was assessed generated ten categories, showing what the articles assessed. These were: Assesses skills (C01); Assesses assessment strategies, tests, exams and items (C02); Assesses teaching proposals (C03); Assesses an instrument, tool or technology (C04); Assesses understandings, perceptions and interest (C05); Assesses knowledge (C06); Assesses the curriculum (C07); Assesses attitudes (C08); Assesses other research (C09); and Discusses assessment in different contexts (C10). Among the categories, the assessment of teaching proposals was the most prominent (19%), followed by the assessment of instruments and technologies (17.5%). The perspectives of leading international authors point to the need for a more plural approach to assessments, emphasizing the potential of assessment in collecting data and making inferences about learning. Among the gaps identified are the lack of long-term studies and the need to better understand students' perceptions and the impact

of assessment practices. The results of this research reveal a significant transformation in publications on Assessment in the field of Chemistry Education over the time frame and highlight the need for future research to address existing gaps, especially in relation to the assessment of teaching practices and the context of Basic Education. In future research, the characteristics, trends and gaps presented in this study can serve as a starting point to promote a more comprehensive and well-founded understanding of the topic of assessment in different educational contexts.

Keywords: Assessment. Systematic Literature Review. Chemistry Education. Researchers. International Context.

LISTA DE FIGURAS

Figura 1	– Triângulo da Avaliação	23
Figura 2	– Frequência absoluta de artigos publicados ao longo dos 20 anos	57
Figura 3	– Distribuição dos artigos por continente	65
Figura 4	– Porcentagem de artigos com foco principal em avaliação/avaliar	134

LISTA DE QUADROS

Quadro 1 –	Um guia de oito passos para uma Revisão Bibliográfica Sistemática	40
Quadro 2 –	Cronograma das atividades de pesquisa	41
Quadro 3 –	Resultados da busca nas bases de dados	43
Quadro 4 –	Artigos excluídos nas etapas 3 e 4	43
Quadro 5 –	Modelo de Inventário utilizado	46
Quadro 6 –	Artigos excluídos na etapa 6	47
Quadro 7 –	Recodificação por razão das exclusões	49
Quadro 8 –	Participantes das entrevistas e coleta de dados	52
Quadro 9 –	Perguntas da Entrevista	53
Quadro 10 –	Principais autores que publicaram em Avaliação na área de Ensino de Química	59
Quadro 11 –	Periódicos que os artigos foram publicados	60
Quadro 12 –	Países das publicações	63
Quadro 13 –	Colaborações entre Países	64
Quadro 14 –	Continentes das publicações	64
Quadro 15 –	Níveis de Ensino	65
Quadro 16 –	Instituições com maior quantidade de artigos	67
Quadro 17 –	Instituições com menor quantidade de artigos	68
Quadro 18 –	Resumo das categorizações	71
Quadro 19 –	Definições das categorias	72
Quadro 20 –	Definições utilizadas para realizar as categorizações	74
Quadro 21 –	C01: Avalia habilidades	76
Quadro 22 –	C02: Avalia estratégias de avaliação, provas, exames e itens	81
Quadro 23 –	C03: Avalia propostas de ensino	87
Quadro 24 –	C04: Avalia um instrumento, ferramenta ou tecnologia	94
Quadro 25 –	C05: Avalia compreensões, percepções e interesse	103
Quadro 26 –	C06: Avalia conhecimentos	109
Quadro 27 –	C07: Avalia o currículo	112
Quadro 28 –	C08: Avalia atitudes	118
Quadro 29 –	C09: Avalia outras pesquisas	124
Quadro 30 –	C10: Discute avaliação em contextos diferentes	126
Quadro 31 –	Artigos com foco principal em avaliar/avaliação	132
Quadro 32 –	Quantidade de artigos com foco principal em avaliar/avaliação	134
Quadro 33 –	Artigos do primeiro grupo	135

LISTA DE ABREVIATURAS E SIGLAS

ABIS	Acid–Base Interest Scale
ACS	American Chemical Society
ASCI	Attitude toward the Subject of Chemistry Inventory
BEAR	Berkeley Evaluation & Assessment Research Assessment System
BNCC	Base Nacional Comum Curricular
CB-POE	Computer based predict observe explain
CCSSO	Council of Chief State School Officers
CD-ROM	Compact Disc Read-Only Memory
CEPS	Center for Educational Policy Studies
DCARS	Diagnostic Criteria for Anxiety Rating Scale
DOT	Danczak–Overton–Thompson Chemistry Critical Thinking Test
EHEA	European Higher Education Area
ENEM	Exame Nacional do Ensino Médio
ENPEC	Encontro Nacional de Pesquisa em Educação em Ciências
EQulP	Educators Evaluating the Quality of Instructional Products
ERIC	Education Resources Information Center
ERIH Plus	European Reference Index for the Humanities
EUA	Estados Unidos da América
FAST	Formative Assessment for Teachers and Students
IF-ATs	Immediate feedback assessment technique
INEP	Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira
ISCED	International Standard Classification of Education
IUPAC	Union of Pure and Applied Chemistry
JOTSE	Journal of Technology and Science Education
K12	Kindergarten through 12th grade
KS3	Key Stage 3
LTAS	Learning and Teaching Academic Standards
MUST	Math-Up Skills Test
NGSS	Next Generation Science Standards
NRC	National Research Council

PCK	Pedagogical content knowledge
PECEM	Programa de Pós-Graduação em Ensino de Ciências e Educação Matemática
PIBID	Programa Institucional de Bolsas de Iniciação à Docência
PISA	Programme for International Student Assessment
POGIL	Process Oriented Guided Inquiry Learning
RASBQ	Reuniões Anuais da Sociedade Brasileira de Química
SCASS	State Collaborative on Assessment and Student Standards
TCI	Thermochemistry Concept Inventory
TOSRA	Test of Science Related Attitudes
UDL	Universal Design for Learning
UEL	Universidade Estadual de Londrina
UNESCO	United Nations Educational, Scientific and Cultural Organization
WASI	Whimbey Analytical Skills Inventory
WGTCAS	Watson-Glaser Test of Critical Thinking Appraisal
WTL	Writing to Learn

SUMÁRIO

	APRESENTAÇÃO	11
1	INTRODUÇÃO	14
2	FUNDAMENTAÇÃO TEÓRICA	21
2.1	AVALIAÇÃO NO CONTEXTO INTERNACIONAL	21
2.2	AVALIAÇÃO NO CONTEXTO NACIONAL	27
2.3	AVALIAÇÃO SOMATIVA, FORMATIVA E AUTORREGULAÇÃO	31
3	ASPECTOS METODOLÓGICOS	38
3.1	REVISÃO BIBLIOGRÁFICA SISTEMÁTICA	38
3.2	CONTEXTO DA PESQUISA	41
3.3	ANÁLISE DE CONTEÚDO	50
3.4	ENCAMINHAMENTOS METODOLÓGICOS E ANALÍTICOS DAS ENTREVISTAS	52
4	RESULTADOS E DISCUSSÃO	56
4.1	QUAIS SÃO AS PRINCIPAIS CARACTERÍSTICAS DOS ARTIGOS INTERNACIONAIS SOBRE AVALIAÇÃO NA ÁREA DE ENSINO DE QUÍMICA?	56
4.1.1	Uma síntese crítica das tendências e lacunas – Parte I	69
4.2	O QUE TEM SIDO AVALIADO NOS ARTIGOS INTERNACIONAIS SOBRE AVALIAÇÃO NA ÁREA DE ENSINO DE QUÍMICA?	70
4.2.1	Avalia habilidades (C01)	76
4.2.2	Avalia estratégias de avaliação, provas, exames e itens (C02)	81
4.2.3	Avalia propostas de ensino (C03)	87
4.2.4	Avalia um instrumento, ferramenta ou tecnologia (C04)	94
4.2.5	Avalia compreensões, percepções e interesse (C05)	103
4.2.6	Avalia conhecimentos (C06)	109
4.2.7	Avalia o currículo (C07)	112

4.2.8	Avalia atitudes (C08)	118
4.2.9	Avalia outras pesquisas (C09)	124
4.2.10	Discute avaliação em contextos diferentes (C10)	126
4.3	QUAIS ARTIGOS TÊM O FOCO PRINCIPAL EM AVALIAR/AVALIAÇÃO	132
4.4	UMA SÍNTESE CRÍTICA DAS TENDÊNCIAS E LACUNAS - PARTE II	139
4.5	UM RETORNO AO TRIÂNGULO DA AVALIAÇÃO	141
4.5.1	Habilidades e o Triângulo da Avaliação	142
4.5.2	Estratégias de avaliação, provas, exames e itens e o Triângulo da Avaliação	143
4.5.3	Propostas de Ensino e o Triângulo da Avaliação	144
4.5.4	Instrumentos, ferramentas e tecnologias e o Triângulo da Avaliação	145
4.5.5	Compreensões, percepções e interesse e o Triângulo da Avaliação	146
4.5.6	Conhecimentos e o Triângulo da Avaliação	147
4.5.7	Currículo e o Triângulo da Avaliação	149
4.5.8	Atitudes e o Triângulo da Avaliação	150
4.6	ENTREVISTAS COM OS PRINCIPAIS AUTORES INTERNACIONAIS	152
5	CONSIDERAÇÕES FINAIS	168
	REFERÊNCIAS	174
	APÊNDICES	
	APÊNDICE A - Codificação dos 205 artigos revistos nesse estudo	197

APRESENTAÇÃO

Nessa seção, apresento uma breve descrição da minha trajetória pessoal e acadêmica até o momento.

Nasci no Brasil, mais especificamente em Ortigueira, uma pequena cidade agrícola no interior do Paraná. Aos 2 anos de idade, minha vida assumiu um rumo completamente diferente quando minha família e eu nos mudamos para os Estados Unidos. Lá vivi toda a minha infância e uma parte da adolescência em uma península do estado de Massachusetts chamada Cape Cod.

Fui muito privilegiado, pois julgo ter tido uma educação de alta qualidade, inserido em um dos estados que possui uma história de valorizar a educação nos Estados Unidos. Entretanto, nossa intenção nunca foi permanecer no país eternamente, pois éramos imigrantes brasileiros em um país estrangeiro. Quando minha família finalmente decidiu retornar ao Brasil, eu já estava no Ensino Médio. A chegada ao Brasil foi um grande choque, uma mistura de reações positivas e negativas frente a essa nova realidade.

No Ensino Médio, minha disciplina preferida era Química. Eu sentia que tinha uma facilidade natural com os conceitos e o raciocínio utilizado nesta disciplina. Tudo fazia sentido de uma maneira natural, sem esforço. Isso, aliado à minha afinidade com as Ciências Naturais, que cultivava desde pequeno, me fez considerá-la como uma escolha profissional.

Quando optei por fazer o curso de Química no final do Ensino Médio, minhas intenções eram me tornar um Químico. Na época, eu rejeitei a possibilidade de fazer Letras Inglês, mesmo dominando a língua e tendo afinidade, pois tinha medo de acabar me tornando professor. É curioso pensar como minha concepção de “ser professor” era limitada naquela época. E como a minha imaturidade e concepção do mundo me restringia.

Na época, minha família não entendeu o porquê da minha escolha em fazer um curso “atípico” como Química. Todos na minha família eram agricultores, tratoristas, mecânicos ou professores e havia uma expectativa de que eu fizesse Agronomia. Refletindo a respeito desse período, encontro uma ironia hilária e pontual na família. O fato de meu avô paterno se chamar Lavoisier, ou seja, homônimo do *Lavoisier*, o famoso químico francês do século XVIII, considerado o “o pai da Química Moderna”.

Ao ingressar no curso, não posso deixar de mencionar que tive muitas dificuldades. De fato, o curso de Química na UEL é reconhecido por seus egressos como um curso academicamente árduo. Hoje sei que, por ter sido alfabetizado em Inglês e ter a língua inglesa como minha língua nativa, isso não facilitou a experiência em um curso que já tem a fama de ser difícil. Foi um real teste de resiliência, mas após ter concluído, vejo como a experiência da graduação me transformou. Tirei lições valiosas como o poder da disciplina, a importância de não desistir, mesmo quando não parece haver esperança, e o real significado da palavra estudar.

Foi nessa fase que observei o trabalho do professor pesquisador brasileiro e nasceu em mim a vontade de me tornar um também. No geral, fui cativado pela pesquisa e, mais especificamente, pelo Ensino de Química como uma área de pesquisa, mediante o contato com o programa PIBID² e o programa Residência Pedagógica³. Nesse período, tive meus primeiros contatos com a pesquisa na área educacional e senti uma afinidade.

Quando ingressei no Mestrado, tive acesso a um mundo excepcionalmente instigante e estimulante: o Programa de Pós-Graduação em Ensino de Ciências e Educação Matemática (PECEM) da UEL. Fiquei encantado com o comprometimento do quadro docente e com a qualidade do ensino e das pesquisas produzidas no programa. Ali me senti realmente interessado em aprender tudo que eu podia e explorar as inúmeras possibilidades de pesquisa. Decidi entregar o meu máximo e me dedicar à área acadêmica. Não só produzi a minha dissertação sobre Práticas Científicas no Ensino de Ciências (Costa, 2021), mas também publiquei diversos artigos derivados da minha dissertação (Costa; Broietti, 2023; Costa; Broietti, 2022; Costa; Broietti, 2021b; Costa; Broietti, 2021c) e outros artigos sobre temas concomitantes do meu interesse, como Pensamento Crítico, Formação de Professores e Práticas Científicas em outros contextos (Costa *et al.*, 2021a; Cabral *et al.*, 2021; Costa; Broietti, 2021a; Costa *et al.*, 2021b; Costa *et al.*, 2020a; Costa *et al.*,

² O PIBID (Programa Institucional de Bolsa de Iniciação à Docência) é um programa do Ministério da Educação do Brasil, criado em 2007. Seu principal objetivo é promover a formação de professores, oferecendo bolsas a estudantes de licenciatura que atuam em escolas de educação básica. O PIBID busca integrar a teoria à prática, contribuindo para a formação de futuros educadores por meio de experiências em sala de aula.

³ A Residência Pedagógica foi criada em 2016, como parte das políticas do Ministério da Educação do Brasil para fortalecer a formação inicial de professores. O programa tem como foco a articulação entre a teoria e a prática, oferecendo experiências significativas em escolas de educação básica. O objetivo da Residência Pedagógica é promover uma formação mais integrada entre teoria e prática, permitindo que os futuros educadores vivenciem a realidade escolar, desenvolvam projetos pedagógicos sob a supervisão de professores experientes e reflitam sobre sua prática.

2020b; Costa *et al.*, 2020c; Broietti *et al.*, 2019; Costa; Broietti, 2019).

Ao ingressar no Doutorado, eu sabia que queria continuar pesquisando e disseminando o conhecimento. Em colaboração com a minha orientadora, a professora Fabiele, escolhemos trabalhar com a temática da Avaliação na área de Ensino de Química. A escolha foi ideal para ambos de nós, pois envolvia algo que eu tinha prática: realizar uma revisão bibliográfica sistemática, porém dessa vez com um escopo muito maior. O tema também envolvia algo de que eu gostava: a internacionalização da área de Ensino, devido ao trabalho permitir meu contato com pesquisas do mundo todo sobre Avaliação. Além disso, a professora Fabiele, em seu doutorado, havia investigado sobre esse tema e realizado uma revisão sobre Avaliação na área de Ensino de Química no contexto brasileiro, e ela já havia sido questionada a respeito dos resultados no contexto internacional. Logo, meu trabalho viria para responder a tais questionamentos.

Posso dizer que a experiência do Doutorado tem sido tão gratificante quanto a do Mestrado, porém introduz um diferencial. A possibilidade de assumir projetos de pesquisa mais complexos, ambiciosos e com um *corpus* maior, o contato com pesquisadores de outras instituições internacionais que investigam a temática, as leituras, análises e discussões sobre o tema investigado têm igualmente contribuído para o meu amadurecimento como pesquisador. A jornada até aqui é composta de altos e baixos, mas o que tem sido mais gratificante e recompensador é a aprendizagem que tenho obtido do processo e a transformação em mim decorrente disso. Portanto, estou ansioso para ver o que o futuro aguarda, pois sei que ainda há muito a aprender e muito a ser transformado.

1 INTRODUÇÃO

A avaliação é um componente fundamental do processo educacional, crucial para ter uma visão da compreensão do aluno, orientar decisões instrucionais e até mesmo promover a aprendizagem (Black; Wiliam, 1998). No contexto da Química, características particulares, como, por exemplo: o grande volume de conhecimento teórico abstrato presente, sua relação íntima com a alfabetização científica e a presença de habilidades experimentais, tornam a avaliação ainda mais desafiadora. Isso ocorre devido à complexidade de vários conceitos essenciais da disciplina, presentes em tópicos como: teoria atômica, reações químicas e termodinâmica, tornando a avaliação precisa da compreensão do aluno e suas habilidades um ponto central (Dori; Sasson, 2008).

As avaliações tradicionais, frequentemente associadas a provas de múltipla escolha e questões abertas de respostas curtas, têm sido ferramentas avaliativas predominantes no contexto escolar por muitos anos. Entretanto, essas abordagens de avaliação se concentram apenas em medir o conhecimento factual e o entendimento processual dos alunos, verificando a capacidade de memorização de informações dos alunos e a aplicação ou reaplicação de conceitos em contextos já conhecidos (Bransford; Brown; Cocking, 2000).

Nesse sentido, há uma crescente preocupação de que essas avaliações tradicionais possuem grandes limitações para examinar a complexidade do aprendizado do aluno. Hattie e Timperley (2007), por exemplo, argumentam que as avaliações de múltipla escolha e os testes de resposta curta podem fornecer uma visão geral do conhecimento factual do aluno de maneira rápida, porém não são necessariamente eficazes para avaliar o pensamento crítico⁴, a capacidade de analisar e sintetizar informações, ou a aplicação de princípios químicos em contextos novos e não familiares. Portanto, embora as avaliações tradicionais possam fornecer informações sobre o conhecimento superficial dos alunos, estas carecem ao tentar indicar a capacidade dos alunos de integrar e aplicar conhecimentos de maneira mais complexa e criativa.

⁴ O pensamento crítico é um pensar ético e eficaz em vários contextos e domínios para resolver problemas e tomar decisões sobre o que acreditar ou como agir responsável e sustentavelmente. Ele inclui quatro componentes ou dimensões: capacidades, disposições/attitudes, normas e critérios e conhecimentos (Tenreiro-Vieira; Vieira, 2014; Tenreiro-Vieira; Vieira, 2019).

Assim, a necessidade de explorar abordagens de avaliação mais abrangentes e diversificados torna-se evidente. Abordagens alternativas, como avaliações formativas, projetos baseados em problemas e avaliações que promovam habilidades necessárias para a alfabetização científica, como o pensamento crítico, têm sido cada vez mais recomendadas (Hattie; Timperley, 2007). Além disso, para Gibbs e Simpson (2004), a combinação de diferentes formas de avaliação se apresenta como uma opção para proporcionar uma imagem mais completa e precisa do aprendizado dos alunos, bem como para preparar os estudantes a enfrentarem a diversidade de desafios no futuro.

Alguns exemplos de avaliações alternativas, previamente citadas, são: avaliações baseadas em projetos, onde os alunos aplicam conhecimento químico em problemas reais (Thomas, 2000); avaliações de pares, que incentivam a colaboração e a avaliação crítica entre os alunos (Topping, 2009); e a autoavaliação, que promove a metacognição e a construção da responsabilidade pessoal do aluno com a sua própria aprendizagem (Boud; Falchikov, 2007). Para Biggs e Tang (2011), essas abordagens estão alinhadas com modelos pedagógicos mais atuais, que posicionam o aluno no centro dos processos de ensino e aprendizagem e se preocupam com o desenvolvimento de habilidades e não só a verificação de conceitos.

A diversificação de estratégias de avaliação no Ensino de Química é um tema discutido na literatura, pois essencialmente apoia-se cada vez mais a ideia de que a avaliação não pode ser limitada a um único evento somativo no final de uma unidade ou curso, mas sim um processo (Stiggins, 2005; Gibbs; Simpson, 2004; William, 2011). Nesse contexto, William (2011) argumenta que a avaliação é um processo contínuo, que pode incorporar diversas formas ou ferramentas de avaliação ao longo do percurso educacional. Sugere-se que, em vez de implementar avaliações finais e pontuais, a prática avaliativa deve ser flexível e oferecer feedback constante, permitindo ajustes pontuais tanto no ensino do professor quanto na aprendizagem do estudante.

A avaliação formativa, que envolve o feedback regular no processo de aprendizagem, tem ganhado destaque em pesquisas nas últimas duas décadas, devido à sua capacidade de informar não só professores, mas os próprios alunos a respeito do seu progresso. De acordo com Black e William (1998), idealizadores dessa abordagem avaliativa, a avaliação formativa permite intervenções e ajustes ao longo

dos processos de aprendizagem. Para Nicol e Macfarlane-Dick (2006), esse ciclo de feedback em tempo real da avaliação formativa também contribui para a construção de um ambiente onde os alunos podem refletir sobre sua compreensão, identificar áreas para melhoria e se envolver ativamente em seu processo de aprendizagem.

Nosso interesse em compreender um pouco mais acerca da Avaliação na área de Ensino de Química teve início em uma pesquisa de doutorado desenvolvida por Broietti (2013), na qual a pesquisadora investigou dois processos de avaliação de ingresso no curso de Química da Universidade Estadual de Londrina: o Vestibular e o ENEM⁵. Nessa pesquisa, Broietti (2013) buscou investigar as questões de Química desses exames, segundo quatro critérios: os conhecimentos e conteúdos químicos necessários para sua resolução; as competências/habilidades nelas exigidas; os aspectos contextuais; e os aspectos estruturais contemplados na elaboração dessas questões.

Como resultado, a pesquisadora evidenciou que, embora tenham sido encontradas algumas sobreposições em termos de conhecimentos e conteúdos químicos abordados, competências e habilidades, contextos e representações, nas questões presentes nos dois exames, existem diferenças significativas que implicam finalidades divergentes. A autora também enfatiza as implicações dessas avaliações para o Ensino de Química, salientando que os conhecimentos, temas, habilidades e representações que estão presentes nesses exames e como são exigidos nos processos de avaliação influenciam tendências educativas que nem sempre correspondem aos promissores objetivos da Educação Científica/Química (Broietti, 2013).

Em um artigo originado dessa pesquisa, Broietti, Santin Filho e Passos (2015) apresentam os resultados da análise de artigos nacionais focados em Avaliação em Química, publicados em periódicos da área de Ensino. Embora os autores tenham notado um crescente interesse pelo desenvolvimento de pesquisas relacionadas ao campo da avaliação em Química, como demonstrado pela grande maioria (90%) dos artigos publicados a partir de 2005. Considerando o recorte temporal de 1986 a 2012, de um total de 2.338 artigos, apenas 17 deles (menos de 1%) estavam relacionados à temática de interesse. Além disso, para caracterizar o

⁵ O ENEM, ou Exame Nacional do Ensino Médio, é uma prova realizada anualmente no Brasil. A prova foi criada pelo Ministério da Educação e serve para avaliar o desempenho dos estudantes ao final do ensino médio, bem como serve para ingresso em instituições de ensino superior.

campo da Avaliação em Química, os autores classificaram esses dezessete artigos em três categorias: o uso de instrumentos avaliativos, a análise de questões e/ou programas de avaliação, a avaliação e a formação de professores. O total de artigos em cada categoria foi de 5, 6 e 6, respectivamente.

Seguindo essa mesma linha de investigação, Ferreira e Broietti (2015) fizeram um levantamento nos anais das Reuniões Anuais da Sociedade Brasileira de Química (RASBQ)⁶ de resumos que abordavam a temática avaliação em Química.

Foram analisados os resumos entre os anos de 2005 e 2014 (10 anos), especificamente na seção EDU (Educação) da RASBQ. Para a análise, foram selecionados 68 resumos referentes à avaliação na área de Ensino de Química, representando cerca de 3,67% do total de resumos publicados, um número baixo, considerando que a avaliação é um dos pilares dos processos de ensino e de aprendizagem (Ferreira; Broietti, 2015).

Os resumos foram distribuídos por categorias, de acordo com a semelhança dos assuntos abordados. Essas categorias incluíram: 1) uso de instrumentos avaliativos; 2) análise de questões e/ou programas de avaliação; e 3) avaliação e a formação de professores. O percentual de resumos em cada categoria foi, respectivamente, 41,17%; 39,71% e 19,12%.

Em outra investigação sobre a temática da Avaliação em Química, Souza e Broietti (2017) analisaram os trabalhos publicados nas atas do Encontro Nacional de Pesquisa em Educação em Ciências - ENPEC⁷, nas edições correspondentes aos anos de 1997 a 2015.

Do total de trabalhos revisados, apenas 19 abordavam a temática investigada, ou seja, Avaliação em Química. Esses trabalhos foram distribuídos em categorias estabelecidas por Broietti, Santin Filho e Passos (2013), sendo acrescentada uma nova categoria denominada Levantamento Bibliográfico, que incorpora a análise da avaliação em Química em produções acadêmicas, publicações em periódicos e anais/atas de congressos.

⁶ A RASBQ é o maior evento de Química na América Latina, com a participação de aproximadamente 2500 colaboradores. São apresentadas cerca de 2200 comunicações em cada reunião anual, distribuídas nas diversas áreas de atuação.

⁷ O Encontro Nacional de Pesquisa em Educação em Ciências (ENPEC) é um evento bienal promovido pela Associação Brasileira de Pesquisa em Educação em Ciências (ABRAPEC), que tem como objetivo reunir e favorecer a interação entre os pesquisadores das áreas de Educação em Biologia, Física, Química e outras áreas relacionadas, enfocadas individualmente ou de forma interdisciplinar, com o objetivo de discutir trabalhos de pesquisa recentes e tratar de temas de interesse.

Assim como nas demais produções relatadas, observa-se que este é um tema recente na pauta dos pesquisadores nacionais, e o número de publicações a respeito deste tema vem crescendo ao longo dos anos. Dos 19 trabalhos selecionados, 6 foram alocados na categoria de uso de instrumentos avaliativos, evidenciando a importância da avaliação para os processos de ensino e aprendizagem, pois é por meio dela que o professor coleta informações sobre a aprendizagem dos estudantes e reavalia propostas de ensino.

Outros seis trabalhos foram submetidos à categoria de análise de questões e/ou programas de avaliação, pois examinavam questões relacionadas ao conhecimento químico em exames de larga escala, além de analisarem programas e currículos. Essa análise é relevante para entender como o conhecimento químico está sendo expresso nos exames de larga escala (Souza; Broietti, 2017).

Na categoria avaliação e formação de professores, o número de trabalhos submetidos foi quase igual ao das outras categorias mencionadas, totalizando cinco investigações que abordavam aspectos relativos à avaliação em Química associados à formação de professores. A categoria com o menor número de trabalhos foi a de levantamento bibliográfico, com apenas 2 estudos que analisavam aspectos relacionados à temática Avaliação no Ensino de Química em produções acadêmicas, periódicos e anais/atas de congressos. Esses dados confirmam que o estudo desta temática ainda é reduzido e recente no Brasil e, consequentemente, existem poucos trabalhos publicados (Souza; Broietti, 2017).

Das discussões originadas dessas pesquisas, constatamos ser necessária uma revisão similar no contexto internacional, devido às indagações e questionamentos dos membros da banca de doutorado da pesquisa mencionada e dos referees dos trabalhos. Além disso, tínhamos a intenção de investigar se essa mesma tendência, em termos de características e lacunas nos estudos sobre Avaliação na área de Ensino de Química, se repetia no contexto internacional.

Assim, nesta investigação, nosso objetivo foi:

- I) Identificar e analisar as principais características, tendências e lacunas dos artigos internacionais sobre Avaliação na área de Ensino de Química;
- II) Identificar e analisar o que tem sido avaliado nos artigos internacionais sobre Avaliação na área de Ensino de Química;

III) Identificar e analisar as perspectivas dos principais autores internacionais em Avaliação na área de Ensino de Química sobre as características, tendências e as lacunas das pesquisas.

As questões que se colocaram, portanto, foram:

- I) Quais são as principais características, tendências e lacunas dos artigos internacionais sobre Avaliação na área de Ensino de Química?
- II) O que tem sido avaliado nos artigos internacionais sobre Avaliação na área de Ensino de Química?
- III) Quais são as perspectivas dos principais autores internacionais em Avaliação na área de Ensino de Química sobre as características, tendências e as lacunas das pesquisas?

A partir do exposto, ressaltamos que a presente tese não busca tratar especificamente da avaliação da aprendizagem, da avaliação de desempenho, da avaliação institucional ou da avaliação em larga escala, mas sim analisar pesquisas em que o tema avaliação é o objeto de investigação. Nesse sentido, a produção bibliográfica investigada foi constituída por artigos da área de Ensino de Química publicados em periódicos internacionais, sendo considerados os artigos de duas bases de dados (ERIC e ERIH Plus), publicados em um período de vinte anos (2002-2021).

Essa tese foi organizada em cinco capítulos, seguido das referências. Na sequência, descrevemos cada capítulo:

No Capítulo 1 apresentamos a Introdução, situando brevemente a temática Avaliação e apresentando as questões de pesquisa e os objetivos dessa tese.

No Capítulo 2, intitulado Fundamentação Teórica, apresentamos os conceitos de avaliação conforme os principais referenciais internacionais e nacionais, destacando especialmente a Avaliação Formativa, contrastando-a com a Avaliação Somativa, dada a crescente discussão sobre o tema na literatura recente.

No Capítulo 3 – Aspectos Metodológicos – apresentamos os fundamentos da revisão bibliográfica sistemática e descrevemos as oito etapas para a realização de uma revisão bibliográfica sistemática conforme Okoli (2015) e a

realização dessas etapas nessa pesquisa. Na sequência, explicitamos a metodologia de análise de dados utilizada, a Análise de Conteúdo, e descrevemos como os movimentos analíticos foram realizados para essa pesquisa. Por fim, apresentamos os encaminhamentos metodológicos e analíticos realizados para entrevistar os principais autores em Avaliação na área de Ensino de Química sobre as características, tendências e as lacunas das pesquisas.

No Capítulo 4 – Resultados e Discussão – separamos em quatro partes que buscam responder cada uma das quatro perguntas de pesquisa individualmente. Nas seções deste capítulo, discorremos sobre: as características das publicações; o que tem sido avaliado nas publicações; as tendências e lacunas de pesquisa; e as perspectivas dos principais autores em Avaliação na área de Ensino de Química.

Por fim, no Capítulo 5 – Considerações finais – apresentamos as compreensões e interpretações conclusivas, indicando implicações para estudos futuros.

Nas referências, incluímos a bibliografia utilizada na elaboração desta tese.

2 FUNDAMENTAÇÃO TEÓRICA

Neste capítulo apresentamos o conceito de avaliação segundo alguns dos principais referenciais assumidos no contexto internacional, considerando o *corpus* desta pesquisa. Também apresentamos o conceito de avaliação conforme referenciais nacionais. Finalmente, apresentamos o conceito de Avaliação Formativa e descrevemos suas principais características, contrastando-a com a Avaliação Somativa, considerando que a primeira tem sido muito discutida na literatura em Avaliação dos últimos anos.

2.1 AVALIAÇÃO NO CONTEXTO INTERNACIONAL

Para Stowe e Cooper (2019):

Saber o que os alunos sabem e conseguem fazer é desafiador, pois não podemos examinar diretamente o que qualquer pessoa está pensando. Em vez disso, devemos construir oportunidades para que os alunos tornem seu pensamento visível para que possamos interpretar as evidências extraídas dessas oportunidades para inferir a respeito do progresso em direção aos resultados desejados (Stowe; Cooper, 2019, p. 1, tradução nossa)⁸.

A citação expressa uma das grandes dificuldades encontradas na avaliação: a impossibilidade de saber, exatamente, o que um sujeito sabe sobre determinado assunto, conceito ou tema. Portanto, somos obrigados a fazer inferências a partir do que o sujeito fala que sabe ou faz. Na sala de aula, isso é tradicionalmente realizado por meio do desempenho do sujeito em itens de avaliação, mas não está limitado apenas a isso.

A noção de avaliação como “um processo de raciocínio a partir de evidências” geradas por alunos em tarefas de avaliação tem ganhado ampla aceitação na comunidade de Ensino de Ciências no contexto internacional (Stowe; Cooper, 2019). O relatório/documento *Knowing What Students Know: The Science and Design of Educational Assessment* do *National Research Council* (NRC, 2001) traz essa ideia.

O NRC (2001) é um documento importante no contexto internacional, pois, diferente de relatórios anteriores do NRC, este fornece parâmetros norteadores

⁸ No texto original: “Knowing what students know and can do is challenging business, as we cannot directly examine what any given person is thinking. Instead, we must construct opportunities for students to make their thinking visible and interpret the evidence elicited from these opportunities to infer progress toward desired outcomes” (Stowe; Cooper, 2019, p. 1).

educacionais que colocam a avaliação em foco. Nele, algumas das questões mais urgentes e complexas em avaliação educacional são exploradas.

A respeito do foco do relatório, o NRC (2001) discute:

[O relatório] examinou o papel e os usos apropriados da avaliação na reforma baseada em padrões, um movimento que está remodelando a educação em todo o país. No processo de realização deste trabalho, o conselho e seus comitês se aprofundaram em questões fundamentais sobre avaliação educacional, como quais são seus propósitos; quais tipos de conhecimento e habilidades devem ser avaliados; quão bem as avaliações atuais, como a Avaliação Nacional do Progresso Educacional, estão atendendo às várias demandas colocadas sobre elas; e quais novos desenvolvimentos prometem melhorar a avaliação (NRC, 2001, p. 6, tradução nossa)⁹.

Para construir o relatório, um comitê de especialistas/pesquisadores foi proposto pelo NRC para revisar os avanços nas ciências cognitivas e nas ciências de medição/avaliação, para considerar as implicações na remodelação da avaliação educacional. Assim, essa reformulação ganha confiabilidade por ser baseada em evidências e resultados de pesquisas da área científica educacional. Isso é importante, pois frequentemente escutamos queixas por parte da equipe de professores na Educação Básica, de que os parâmetros educacionais estabelecidos carecem de compreensão a respeito da realidade escolar e não consideram os resultados de pesquisas já realizadas na área.

Os tópicos abordados na confecção do NRC (2001) são:

(1) práticas de avaliação baseadas em princípios cognitivos e que estão sendo implementadas com sucesso em escolas e salas de aula, (2) novos modelos estatísticos promissores para uso na avaliação de uma ampla gama de desempenhos cognitivos, (3) programas que envolvem os alunos na autoavaliação e na avaliação por pares, (4) tecnologias inovadoras para aprendizagem e avaliação, (5) programas de intervenção instrucional baseados cognitivamente e (6) perspectivas políticas sobre novas formas de avaliação (NRC, 2001, p. 7, tradução nossa)¹⁰.

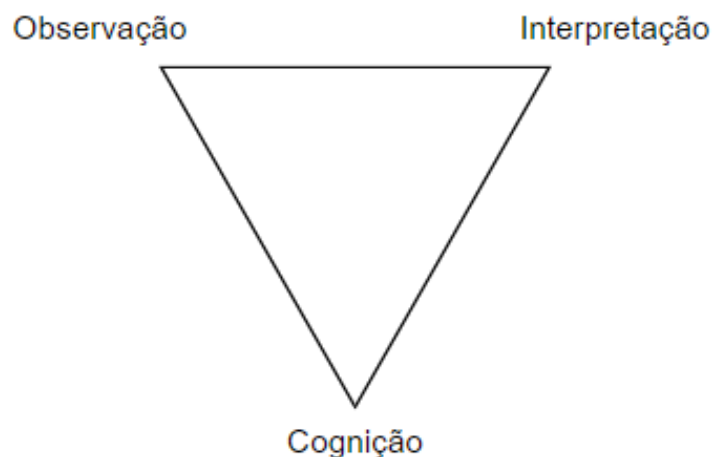
⁹ No texto original: “[The report] examined the role and appropriate uses of assessment in standards-based reform, a movement that is reshaping education throughout the country. In the process of carrying out this work, the board and its committees have delved into fundamental questions about educational assessment, such as what its purposes are; which kinds of knowledge and skills should be assessed; how well current assessments, such as the National Assessment of Educational Progress, are fulfilling the various demands placed on them; and which new developments hold promise for improving assessment” (NRC, 2001, p. 6).

¹⁰ No texto original: “(1) assessment practices that are based on cognitive principles and are being successfully implemented in schools and classrooms, (2) new statistical models with promise for use in assessing a broad range of cognitive performances, (3) programs that engage students in self- and peer assessment, (4) innovative technologies for learning and assessment, (5) cognitively based instructional intervention programs, and (6) policy perspectives on new forms of assessment” (NRC, 2001, p. 7).

Para o NRC (2001), embora as avaliações possam variar muito dependendo do contexto e do propósito, estas compartilham alguns princípios fundamentais, como sua característica de ser inerentemente um processo de raciocínio baseado em evidências. Além disso, é preciso considerar que, independentemente da estratégia de avaliação escolhida pelo professor, nele sempre haverá um nível de imprecisão. Isso ocorre, pois a avaliação não fornece o conhecimento do sujeito, mas sim estimativas do conhecimento e das habilidades de um sujeito (NRC, 2001).

Assumindo essa característica inerente às avaliações, o NRC (2001) apresenta a abordagem da “avaliação como um argumento evidencial”, representada na Figura 1.

Figura 1 – Triângulo da Avaliação



Fonte: extraído e adaptado de NRC (2001, tradução nossa)

Na Figura 1 vemos três facetas interligadas nos cantos de um “triângulo da avaliação”. Nessa abordagem, o elaborador de uma avaliação deve considerar três pontos: (1) um modelo de cognição e aprendizagem do aluno em um determinado domínio; (2) os tipos de observações que fornecerão evidências das competências dos alunos; (3) um processo de interpretação para dar sentido às evidências (NRC, 2001). Nota-se que os três elementos do triângulo não existem isoladamente, mas estão conectados e são dependentes uns dos outros. Esses três elementos: cognição, observação e interpretação, precisam ser integrados de maneira

consciente e cuidadosa para garantir a maior precisão possível da avaliação e uma utilidade de suas conclusões.

O vértice cognitivo do triângulo da avaliação se refere a uma teoria ou conjunto de crenças sobre como os alunos representam o conhecimento e desenvolvem competência em um domínio de assunto. Em qualquer avaliação, uma teoria de aprendizagem no domínio é necessária para identificar o conjunto de conhecimentos e habilidades que é importante avaliar para a tarefa em questão (NRC, 2001).

Para o NRC (2001), a cognição se refere aos processos mentais envolvidos na aquisição e compreensão do conhecimento por meio do pensamento, experiência e sentidos. A cognição inclui várias habilidades, como percepção, memória, raciocínio e resolução de problemas. Em contextos educacionais, a cognição é particularmente importante para entender como os alunos processam informações e desenvolvem habilidades. Aspectos-chave da cognição incluem estruturas de conhecimento, que consistem em esquemas. Esses esquemas são estruturas mentais que ajudam os alunos a organizar e interpretar informações. Por exemplo, os alunos constroem esquemas para conceitos matemáticos que lhes permitem conectar novas informações com o que já sabem (NRC, 2001). A compreensão conceitual vai além da mera recordação factual, envolvendo a compreensão de conceitos em um nível profundo e a aplicação desses conceitos em diferentes contextos (Bransford; Brown; Cocking, 2000).

Para Stowe e Cooper (2019), o vértice cognitivo do triângulo da avaliação envolve uma teoria baseada em evidências de como os alunos desenvolvem, organizam e usam o conhecimento em um domínio. Essa teoria pode modelar uma progressão de aprendizagem em uma disciplina. As progressões de aprendizagem são um dos “modelos de cognição” mais aceitos pela literatura atualmente, indicando tarefas para facilitar a distinção de uma compreensão cada vez mais elevada.

É importante pontuar que as formas sucessivamente mais complexas de pensar, apresentadas por uma teoria da cognição não implicam que todos os alunos se movem pela mesma progressão de maneira linear ou igual, dado que há uma pluralidade de caminhos que podem ser seguidos para construir conhecimento útil de um domínio. As progressões de aprendizagem são apenas uma ferramenta

para facilitar a identificação da progressão em um tema, e consequentemente a distinção entre esses níveis.

Em relação às implicações da cognição para avaliação, o NRC (2001) discute que é importante que as avaliações se alinhem com os objetivos de aprendizagem formulados para o contexto. Estas necessitam refletir com precisão os processos cognitivos envolvidos na aprendizagem do conteúdo específico para evidenciar o que os alunos realmente devem saber e fazer. Para isso, a utilização de diversos tipos de avaliação, como avaliações formativas, somativas, autorreguladoras e baseadas em desempenho, permitem uma compreensão mais abrangente da cognição do aluno (NRC, 2001; Black; Wiliam, 1998). Por fim, avaliações eficazes também devem oferecer feedback significativo para os alunos, para que estes possam refletir sobre seus processos cognitivos e melhorar suas estratégias de aprendizagem (NRC, 2001).

A respeito do vértice da observação do triângulo da avaliação:

O canto de observação do triângulo de avaliação representa uma descrição ou conjunto de especificações para tarefas de avaliação que irão provocar respostas esclarecedoras dos alunos. Em uma sessão de tutoria, por exemplo, a estrutura de observação descreve o que o aluno diz e faz, não diz e faz, ou diz ou faz com tipos específicos de suporte [...] (NRC, 2001, p. 48, tradução nossa)¹¹.

Assim, esse vértice representa as observações consideradas convincentes a respeito da compreensão do aluno. O vértice da observação também representa as tarefas de avaliação capazes de provocar essas observações, podendo ser respostas escritas ou orais, ou até mesmo a escolha de uma alternativa incorreta dentre itens de múltipla escolha, por exemplo.

A respeito do terceiro vértice do triângulo da avaliação, o NRC (2001) comenta que o vértice da interpretação:

[...] abrange todos os métodos e ferramentas usados para raciocinar a partir de observações falíveis. Ele expressa como as observações derivadas de um conjunto de tarefas de avaliação constituem evidências sobre o conhecimento e as habilidades que estão sendo avaliados (NRC, 2001, p. 48, tradução nossa)¹².

¹¹ No texto original: "The observation corner of the assessment triangle represents a description or set of specifications for assessment tasks that will elicit illuminating responses from students. In a tutoring session, for example, the observation framework describes what the learner says and does, does not say and do or says or does with specific kinds of support [...]" (NRC, 2001, p. 48).

¹² No texto original: "The interpretation corner of the triangle encompasses all the methods and tools used to reason from fallible observations. It expresses how the observations derived from a set of assessment tasks constitute evidence about the knowledge and skills being assessed" (NRC, 2001, p. 48).

Aqui vemos quão intimamente ligado está o vértice da interpretação e o vértice da observação, já que as observações é que são interpretadas. O NRC (2001) discute que o triângulo da avaliação pode ser empregado tanto no contexto de avaliações realizadas em sala de aula como em avaliações em larga escala¹³. No contexto de avaliação em larga escala, a interpretação é geralmente composta por um modelo estatístico, o qual é uma caracterização ou resumo de padrões que se esperaria ver nos níveis variados de competência do aluno.

O triângulo da avaliação se caracteriza como uma estruturação conceitual útil para melhor compreender diferentes aspectos de uma avaliação e garantir que esta seja realizada de uma forma abrangente, seja do desempenho, compreensão, atitude, etc. A respeito das atitudes, estas impactam todos os aspectos da avaliação. Atitudes positivas, como motivação e confiança podem melhorar o desempenho cognitivo e o engajamento dos alunos, enquanto atitudes negativas como ansiedade ou um baixo nível de interesse podem prejudicá-las. Conforme o triângulo de avaliação as atitudes podem ser consideradas nos três vértices. No vértice da cognição, atitudes positivas/negativas promovem/prejudicam a aprendizagem; no vértice da observação, as atitudes influenciam como os alunos desempenham suas tarefas, afetando a forma como são observadas; no vértice da interpretação, os resultados da avaliação devem considerar as atitudes, pois elas podem impactar o desempenho e os resultados, como discutido anteriormente.

De acordo com Hadji (2001), em sua obra *Avaliação Desmistificada*, refletir a respeito do papel da avaliação é uma questão central para compreendermos como os alunos podem aprender e se desenvolver. É apresentada uma visão desconstruída da avaliação, muito além de um simples processo de medir, mas como uma ferramenta fundamental para as aprendizagens. Assim, sugere-se que a avaliação esteja sempre a serviço dos processos de ensino e de aprendizagem, ajudando os alunos a entenderem seus próprios caminhos para o aprendizado e crescimento.

Hadji (2001) instiga uma reflexão a respeito da avaliação de maneira mais ampla, indo além de um simples número ou conceito. Nesse sentido, avaliar não

¹³ As avaliações em larga escala são elaboradas por um órgão externo às escolas, com a finalidade de fazer juízos de valor e propor alternativas em âmbito mais amplo que o da instituição de ensino (Brasil, 2010).

é apenas medir, mas sim realizar um confronto que envolve negociação, interpretações e trocas. São discutidas algumas características frequentemente negligenciadas da avaliação, como a subjetividade do avaliador e a impossibilidade de uma avaliação totalmente neutra. Isso devido à natureza do contexto escolar, carregado de emoções e inúmeras influências sociais. Logo, para Hadji (2001) a avaliação nunca pode ser vista como algo puramente objetivo; ela é sempre influenciada pela visão do professor/avaliador e pelas suas experiências. Nesse entendimento o ato de avaliar não pode ser resumido a aplicação de um critério meramente impessoal, mas em um julgamento em que o avaliador interpreta e se posiciona diante da realidade de cada aluno.

A fim de explicitar sua visão da avaliação, Hadji (2001) apresenta três afirmações essenciais: I) a avaliação é subjetiva, pois sempre envolve a intuição e a interpretação do avaliador; II) a avaliação envolve a elaboração de um julgamento sobre o aluno e sua realidade, o que exige uma interpretação cuidadosa de cada situação; e III) a avaliação precisa ser qualitativa e formativa, sendo integrada aos processos de ensino e de aprendizagem, e jamais como um evento isolado ou punitivo.

2.2 AVALIAÇÃO NO CONTEXTO NACIONAL

A avaliação no contexto brasileiro tem passado por transformações significativas nas últimas décadas, refletindo mudanças nas concepções de ensino e aprendizagem. De acordo com Pereira (2018), historicamente, a avaliação foi vista predominantemente como um momento de verificação da aprendizagem, frequentemente reduzida à atribuição de uma nota que sintetizava o desempenho dos alunos. Apesar disso, a forma como a comunidade docente e de pesquisa têm visto a avaliação tem se modificado à medida que se reconhece que ela é um processo contínuo e formativo, que não apenas mensura o desempenho dos alunos, mas também orienta o ensino e promove a melhoria da prática pedagógica.

No Brasil, a avaliação educacional é regulamentada por diversas políticas públicas, como a Lei de Diretrizes e Bases da Educação Nacional (LDB, 1996) e as orientações do INEP (2020). A LDB destaca a importância da avaliação para o desenvolvimento dos alunos, ressaltando a necessidade de que esta seja inclusiva e contextualizada. A ênfase na inclusão é particularmente relevante, uma

vez que busca garantir que todos os alunos, independentemente de suas condições sociais, econômicas ou culturais, tenham acesso a uma educação de qualidade. Isso implica que as práticas avaliativas devem ser adaptadas para atender às diversas realidades dos estudantes, promovendo um ambiente de aprendizagem mais equitativo.

Por outro lado, avaliações em larga escala, como o Exame Nacional do Ensino Médio (ENEM), servem a múltiplos propósitos: além de permitir o acesso a instituições de ensino superior, elas também diagnosticam a qualidade da educação em diferentes regiões do país e podem ser usadas para informar políticas públicas educacionais (INEP, 2020). A utilização do ENEM como ferramenta para a formulação de políticas educacionais destaca a importância da avaliação não apenas como um instrumento de mensuração, mas também como um agente importante para promover mudanças estruturais no sistema educacional.

No que se refere ao Ensino de Ciências, a avaliação assume um papel crucial. A Base Nacional Comum Curricular (BNCC) destaca a importância da avaliação no desenvolvimento de competências e habilidades relacionadas ao conhecimento científico, indo além da mera memorização de conteúdos (Brasil, 2017). Nesse sentido, a avaliação precisa considerar não apenas a compreensão conceitual, mas também a capacidade dos alunos de aplicar o conhecimento em situações práticas e cotidianas. Essa abordagem integrada é fundamental para preparar os alunos para enfrentarem os desafios do momento atual, que exigem habilidades de pensamento crítico, resolução de problemas e criatividade.

A ideia de que a avaliação está integrada às práticas pedagógicas para promover a melhoria contínua também tem ganhado destaque na literatura educacional. Fernandes (2020) argumenta que a avaliação não existe como um processo isolado, mas sim como parte integrante do currículo. Essa integração é importante, ao permitir que professores ajustem suas abordagens de ensino com base nas necessidades identificadas por meio de instrumentos avaliativos. Além disso, a avaliação fornece oportunidades para a reflexão crítica tanto por parte dos educadores quanto dos alunos, promovendo um ciclo de aprendizado contínuo.

Fernandes (2020) discute cinco ideias-chave sobre avaliação pedagógica que facilitam sua compreensão como algo que vai além da mera classificação. Essas ideias incluem: 1) a avaliação precisa contribuir para a aprendizagem dos alunos; 2) a avaliação é um processo pedagógico, não meramente

técnico; 3) a avaliação deve focar na melhoria do ensino, não apenas na classificação; 4) a avaliação é subjetiva, mas pode gerar resultados credíveis; e 5) o feedback é central na avaliação, orientando os alunos sobre onde podem melhorar. Esse enfoque no feedback é essencial, ao oferecer aos alunos orientações claras sobre seu desempenho, possibilitando um aprendizado mais autônomo e direcionado.

Uma mudança paradigmática importante na própria conceituação da avaliação tem sido o abandono da visão puramente classificatória para a ideia de apropriação de uma perspectiva que valoriza a construção de significados (Silva; Moradillo, 2002). Essa abordagem permite que a avaliação seja vista como um processo colaborativo, no qual professores e alunos trabalham juntos para construir conhecimento. Nesse sentido, argumenta-se que a avaliação seja contínua, permitindo que professores verifiquem a apropriação dos conteúdos e os alunos analisem sua própria aprendizagem.

A avaliação é, portanto, um processo complexo que deve ser adaptado ao contexto de cada sala de aula. Esse reconhecimento da diversidade nas práticas avaliativas é essencial para garantir que todos os alunos tenham a oportunidade de demonstrar seu aprendizado de formas que sejam significativas para eles. Nesse sentido, a formação continuada de professores é crucial, ao oferecer a oportunidade de discutir e refletir a respeito da problemática que ocorre no cotidiano da sala de aula e a respeito das diferentes abordagens avaliativas que possam ser implementadas em suas práticas docentes diárias.

Fernandes (2011) discute conceitos estruturantes na avaliação pedagógica, destacando a necessidade de uma transformação nas práticas educacionais. As escolas enfrentam desafios que exigem mudanças de paradigmas, distantes das concepções tradicionais do século passado. Assim, é fundamental que os professores atuem como sujeitos reflexivos, capazes de reinventar o currículo e integrar a avaliação nos processos de ensino e aprendizagem. Essa transformação não é apenas necessária para melhorar a qualidade da educação, mas também para preparar os alunos para serem cidadãos críticos e engajados em suas comunidades.

Além disso, Fernandes (2011) ressalta que a avaliação é uma prática social complexa, que necessita ser transparente e compreendida por todos os sujeitos envolvidos. A articulação de diferentes tipos de avaliação, como a formativa e somativa, em uma diversidade de estratégias, é essencial para garantir a credibilidade e eficácia da avaliação pedagógica. Essa diversidade nas abordagens de avaliação

proporciona uma visão mais ampla e integral do aprendizado dos alunos. Essa variedade não se limita apenas ao conhecimento acadêmico, mas também abrange fatores emocionais, sociais e comportamentais que impactam o aprendizado. Além disso, ao reconhecer que cada aluno possui um estilo e ritmo próprios, essa abordagem inclusiva permite que os educadores adaptem suas práticas pedagógicas de forma mais eficaz, atendendo às necessidades individuais e criando um ambiente de aprendizado mais relevante e produtivo.

Os conceitos de avaliação formativa e somativa, introduzidos por Scriven (1967) e expandidos por Bloom, Hastings e Madaus (1971), apontam um marco na avaliação educacional tanto no contexto internacional como no contexto nacional. Essas discussões promovem práticas inovadoras, como a autoavaliação e a diversificação de abordagens de avaliação, ressaltando a importância de considerar não apenas o conhecimento acadêmico, mas também aspectos emocionais e afetivos. A ênfase nestes aspectos é fundamental para atender às diversas necessidades dos alunos, reconhecendo que o aprendizado é um processo multifacetado que envolve diferentes dimensões do ser humano.

Bloom (1956) enfatizou que o ensino deve transcender os níveis básicos de sua taxonomia, promovendo processos de pensamento mais complexos, como análise e síntese. Já durante os anos da década de 70 e 80, a avaliação pedagógica ganhou relevância, com o incentivo a tarefas de resolução de problemas e atividades experimentais (Fernandes, 2011). A evolução da avaliação formativa nos anos 80, que passou a ocorrer durante os processos de ensino e aprendizagem, destacou o feedback como um elemento crucial, devendo ser fornecido de maneira descritiva e frequente para apoiar a melhoria das aprendizagens. Essa mudança reflete uma compreensão mais profunda da aprendizagem como um processo dinâmico, que exige adaptações constantes.

Apesar dessas discussões, ainda persistem confusões sobre a natureza da avaliação formativa, com alguns acreditando que ela é sempre qualitativa e a somativa, quantitativa. Contudo, ambas podem empregar dados qualitativos e quantitativos, dependendo do propósito da informação obtida. Black e Wiliam (2006) e Harlen (2005, 2006) fornecem distinções importantes entre esses tipos de avaliação. Enquanto a avaliação somativa registra o que foi aprendido, a formativa busca influenciar os processos de ensino e aprendizagem. Harlen (2005) argumenta que,

embora conceitualmente distintas, as duas se inter-relacionam na prática, formando uma dimensão contínua.

Em suma, a avaliação pode ser considerada um processo sistemático e contextualizado, essencial para entender as capacidades dos alunos e reestruturar o ensino do professor. Ela vai além da simples atribuição de notas, utilizando observações, conversas informais, provas e relatórios, garantindo a precisão das informações coletadas (Pereira, 2009). Essas informações são fundamentais para guiar decisões pedagógicas e aprimorar os processos de ensino e de aprendizagem (Figueiredo, 2014).

Um dos principais equívocos na comunidade docente é confundir avaliação com classificação. Assim, a avaliação é entendida como uma prática pedagógica que considera aspectos éticos, políticos e didáticos, visando ajudar alunos e professores a aprimorarem seu aprendizado (Santos, 2012). Diante desse cenário, é fundamental investir na formação inicial e continuada de professores, no desenvolvimento de recursos pedagógicos diversificados e na criação de um ambiente escolar que favoreça a inovação e a reflexão crítica. Essa abordagem não apenas enriquece a prática docente, mas também capacita os alunos a se tornarem aprendizes autônomos e críticos, preparados para enfrentar os desafios do mundo contemporâneo.

Na sequência, apresentamos de forma um pouco mais detalhada as características das avaliações formativa, somativa e autorregulação.

2.3 AVALIAÇÃO SOMATIVA, FORMATIVA E AUTORREGULAÇÃO

Segundo Bennett (2011), na Educação Básica, a avaliação formativa apresenta-se claramente como uma tendência nos últimos anos. O tema tem estado presente na pauta de conferências educacionais, discussões a respeito do currículo e em cursos de formação continuada de professores. O conceito de avaliação formativa é comumente contrastado com o de avaliação somativa na literatura. Essa distinção entre características somativas e formativas foi proposta pela primeira vez por Scriven (1967) no contexto de avaliações de programas educacionais (Black e Wiliam 2006; Wiliam e Thompson 2008).

Para Scriven (1967), a avaliação somativa forneceria informações para julgar o valor geral de um programa educacional, enquanto a avaliação formativa

poderia direcionar seus resultados para possibilitar uma melhoria dos programas. Bloom (1969) empregou o conceito de avaliação formativa e avaliação somativa para um olhar centrado nos alunos. Nesse sentido, “o propósito da avaliação formativa era fornecer feedback e corretivos em cada estágio do processo de ensino-aprendizagem” (Bloom, 1969, p. 48, tradução nossa). Já a avaliação somativa era empregada para julgar o que o aluno havia alcançado no final de um curso ou programa. Essa distinção de Bloom (1969) possui relevância, por mais antiga que seja, já que muito do que se construiu a respeito do conceito de avaliação formativa tem sido elaborado a partir de sua conceituação inicial na literatura (Bennett, 2011).

Outro ponto a respeito da distinção entre avaliação formativa e avaliação somativa é a visão de que “avaliação formativa não é uma prova ou um teste, mas um processo” (Popham, 2008, 6, tradução nossa). Logo, a avaliação formativa, compreendida como um processo, vai além da produção de uma pontuação a respeito do aluno, mas um *insight* qualitativo a respeito da sua compreensão (Shepard, 2008). Black e William (1998) ainda discutem que uma das principais características da avaliação formativa é que os resultados são de fato usados para adaptar o ensino para atender às necessidades do aluno, ocorrendo em ciclos curtos entre aulas ou em uma mesma aula.

Uma definição de avaliação formativa publicada pela *Formative Assessment for Teachers and Students (FAST)*, *State Collaborative on Assessment and Student Standards (SCASS)* do *Council of Chief State School Officers (CCSSO)* está contida em McManus (2008):

A avaliação formativa é um processo usado por professores e alunos durante a instrução que fornece feedback para ajustar o ensino e a aprendizagem continuada a fim de melhorar a obtenção dos alunos de resultados instrucionais pretendidos (Mcmanus, 2008, p. 3, tradução nossa)¹⁴.

Apesar da importância das definições de avaliação formativa supracitadas e as discussões a respeito de suas características e funções (Mcmanus, 2008; Black; William, 1998; Popham, 2008; Bloom, 1969; Scriven, 1967), Bennett (2011) faz críticas às simplificações inadequadas que podem ser feitas a respeito da definição de avaliação formativa, caso não haja um cuidado com o termo. De acordo

¹⁴ No texto original: Formative assessment is a process used by teachers and students during instruction that provides feedback to adjust ongoing teaching and learning to improve students' achievement of intended instructional outcomes (Mcmanus, 2008, p. 3).

com Bennett (2011):

É uma simplificação inadequada definir a avaliação formativa apenas como um instrumento porque mesmo o instrumento mais cuidadosamente construído e cientificamente apoiado provavelmente não será eficaz instrucionalmente se o processo que envolve seu uso for falho. Da mesma forma, é uma simplificação inadequada definir a avaliação formativa apenas como um processo, já que mesmo o processo mais cuidadosamente construído provavelmente não funcionará se a “instrumentação”, ou metodologia, sendo usada naquele processo não for bem adequada para o propósito pretendido. “Processo” não pode de alguma forma resgatar instrumentação inadequada, nem a instrumentação pode salvar um processo inadequado. Uma forte conceituação precisa dar atenção cuidadosa a cada componente, bem como a como os dois componentes trabalham juntos para fornecer feedback útil (Bennett, 2011, p. 7, tradução nossa)¹⁵.

Esse cuidado com o termo avaliação formativa é necessário, justamente pelo conceito ter sido altamente utilizado pela comunidade, até de maneira indiscriminada. Uma definição clara desses termos é importante porque, se não conseguimos definir claramente um conceito como esse, não conseguimos documentar sua eficácia, seu impacto e seu uso em pesquisas e/ou intervenções na área.

Assim, a avaliação formativa se caracteriza como fundamental para um ensino mais centrado no aluno. Ao ser entendida como uma parte integral do processo de ensino, a prática avaliativa proporciona feedback qualitativo que permite ajustes nas práticas pedagógicas, atendendo às necessidades específicas dos alunos (McManus, 2008). Desse modo, a avaliação formativa considera as dificuldades e os avanços dos alunos, estabelecendo um ambiente educacional mais inclusivo e eficaz (Black; Wiliam, 1998).

Silva (2017) argumenta que implementar abordagens formativas, considerando todos os pontos discutidos, exige um compromisso significativo dos educadores e uma transformação nas práticas escolares, representando um dos maiores desafios para os sistemas educacionais contemporâneos. Essa mudança é

¹⁵ No texto original: It is an oversimplification to define formative assessment as an instrument because even the most carefully constructed, scientifically supported instrument is unlikely to be effective instructionally if the process surrounding its use is flawed. Similarly, it is an oversimplification to define formative assessment as a process since even the most carefully constructed process is unlikely to work if the ‘instrumentation’, or methodology, being used in that process is not well-suited for the intended purpose. ‘Process’ cannot somehow rescue unsuitable instrumentation, nor can instrumentation save an unsuitable process. A strong conceptualisation needs to give careful attention to each component, as well as to how the two components work together to provide useful feedback (Bennett, 2011, p. 7).

crucial para proporcionar a todos os alunos acesso ao conhecimento e às oportunidades que ele oferece, promovendo sociedades mais justas e equitativas.

Contudo, a discussão sobre o que se entende como “avaliação formativa” e suas características precisa ser realizada com rigor e profundidade. Isso se faz necessário para evitar simplificações ao conceito que podem comprometer sua eficácia, pois, conforme Bennett (2011), a compreensão inadequada desse conceito pode levar a práticas de avaliação que não refletem seu verdadeiro potencial. A clareza conceitual é, portanto, indispensável para garantir que o que se considera como uma “avaliação formativa” em sala de aula cumpra o papel no desenvolvimento educacional, permitindo que professores e alunos se beneficiem de seus princípios (Scriven, 1967; Bloom, 1969).

Hadji (2001) também elabora discussões a respeito do conceito de avaliação formativa, ressaltando sua importância para o desenvolvimento contínuo dos alunos e para a melhoria dos processos de ensino e de aprendizagem. Em seu entendimento, a avaliação formativa não é caracterizada como um momento isolado de medição ou controle, mas como uma ferramenta integrada aos processos de aprendizagem, possuindo como objetivo principal a orientação e o apoio ao aluno em seu desenvolvimento. A avaliação formativa, segundo Hadji (2001), deve ser um processo contínuo de feedback, que permite ao aluno receber informações sobre seu desempenho constantemente, para que este possa refletir a respeito de suas próprias limitações e dificuldades, e buscar formas de superá-los.

Hadji (2001) propõe um conjunto de orientações para tornar a avaliação mais transparente, centrada no aluno, e evidentemente, mais formativa. Para isso sugere-se que o professor não se limite às práticas tradicionais e automáticas de avaliação, como provas tradicionais, mas vá além delas, esclarecendo para os alunos e professores como o processo de aprendizagem ocorre e como a avaliação pode ser uma ferramenta de apoio nesse processo.

Além disso Hadji (2001) sugere que a avaliação empregue critérios claros e compreendidos por todos. A transparência é um ponto chave nesse sentido, pois permite que os alunos saibam exatamente o que é esperado deles e como eles podem melhorar. Para isso, mudanças nas práticas avaliativas dos professores precisam ser realizadas, afastando-as das práticas tradicionais e incorporando práticas mais flexíveis e personalizadas que levam em consideração as particularidades de cada aluno.

Uma última sugestão de Hadji (2001) é a autoavaliação, caracterizada como uma ferramenta valiosa para o autoconhecimento e a autorregulação do aluno. Destaca-se que a autoavaliação não pode ser reduzida a simples reflexão sobre o desempenho; mas precisa desafiar o aluno a realizar uma “autocrítica desestabilizante”, ou seja, o aluno precisa confrontar suas dificuldades e limitações de maneira profunda e construtiva. Hadji (2001) ainda defende que por meio da autoavaliação o aluno pode se tornar mais consciente da sua própria jornada de aprendizagem, aumentando o senso de responsabilidade por ela.

Em resumo, as discussões propostas por Hadji (2001) levam em consideração maneiras de pensar a avaliação de modo a torna-la mais formativa. Nessa perspectiva, a avaliação deixa de ser “um fim em si mesmo”, ou seja, uma simples contagem de acertos e erros, e se reformula como uma prática reflexiva e centrada no aluno. Suas discussões conceituam a avaliação como qualitativa, preocupada com o contexto e as particularidades dos alunos e orientada para o crescimento.

Earl (2012) em *Assessment as learning: Using classroom assessment to maximize student learning*¹⁶ apresenta três abordagens principais para avaliação: avaliação da aprendizagem, avaliação para aprendizagem e avaliação como aprendizagem. As abordagens possuem papéis diferentes, porém essenciais, nos processos educacionais, e por isso recomenda-se que os professores compreendam as nuances de cada uma para maximizar a aprendizagem do aluno.

Para Earl (2012) a avaliação da aprendizagem (avaliação somativa) é a forma mais tradicional de avaliação, sendo frequentemente associada à classificação e avaliações finais. Essa abordagem concentra-se em medir o que os alunos aprenderam no final de um período. Para isso, geralmente utilizam-se avaliações somativas, como exames, projetos finais ou testes padronizados. A avaliação somativa busca resumir o nível de aprendizagem de um estudante em um determinado momento da sua jornada escolar. A avaliação somativa também é utilizada para tomar decisões sobre o desempenho do aluno, a fim de verificar se ele está apto ou não a prosseguir para o próximo nível, conteúdo ou tema de um currículo. Além disso, é por meio da avaliação somativa que os professores geralmente concedem notas para quantificar o desempenho dos alunos.

¹⁶ Avaliação como Aprendizagem: Usando Avaliação em Sala de Aula para Maximizar a Aprendizagem do Aluno.

Na avaliação somativa o foco está nos resultados e geralmente ela é isolada ou separada do processo de aprendizagem. Uma outra limitação da avaliação somativa é a falta de retorno aos alunos a respeito de como melhorar durante o processo de aprendizagem e superar suas limitações naquele determinado momento.

De acordo com Earl (2012) a avaliação para aprendizagem (avaliação formativa) se distingue da anterior, principalmente no sentido de ser mais dinâmica e estar intimamente ligada aos processos de aprendizagem que estão em andamento no momento de sua aplicação. Ao contrário da avaliação somativa, a formativa é focada em reunir informações que ajudam professores e alunos a entender onde os alunos estão em sua jornada de aprendizagem com o objetivo de melhorar os resultados da aprendizagem. Logo o objetivo da avaliação formativa está no âmbito da compreensão/melhoria e não na mensuração/quantificação como na avaliação somativa. Para Earl (2012), utiliza-se o adjetivo *formativa* pois a avaliação ocorre enquanto a aprendizagem ainda está acontecendo, permitindo que ajustes e intervenções sejam realizados para apoiar o progresso dos alunos.

A avaliação formativa pode incluir atividades diversas como: questionários, discussões, observações em sala de aula, feedback dos colegas. Essa pluralidade pode fornecer aos alunos e professores *insights* valiosos sobre o quão bem os alunos estão compreendendo o conteúdo além de ser mais flexível e adaptável às particularidades dos alunos. Earl (2012) ainda discute que a ideia chave da avaliação da aprendizagem (avaliação formativa) é que ela é usada para informar a instrução e orientar os alunos para alcançarem melhores resultados de aprendizagem. Assim, o feedback é central na avaliação formativa, pois é por meio dele que os alunos recebem informações claras para ajudá-los a melhorar.

O terceiro tipo de avaliação discutido por Earl (2012) é a avaliação como aprendizagem (aprendizagem autorregulada). A avaliação como aprendizagem vai além das outras duas, focando em envolver os alunos de forma ativa no processo de avaliação. Na avaliação como aprendizagem, os alunos não são apenas receptores passivos dos resultados da avaliação realizada; mas sim assumem o papel de “avaliadores” de sua própria aprendizagem. Na avaliação como aprendizagem os alunos utilizam atividades para monitorar/refletir a respeito de sua própria aprendizagem para tomarem decisões sobre seu progresso e identificarem áreas onde precisam de melhorias. O ponto chave desse terceiro tipo de avaliação é a autorregulação. Os alunos são encorajados a agir como sujeitos reflexivos e serem

ativos e engajados na avaliação de sua própria compreensão, progresso e estratégias de aprendizagem.

Devido à essas características da avaliação como aprendizagem, Earl (2012) discute que os estudantes desenvolvem habilidades metacognitivas ao se engajarem com esse tipo de avaliação. Nesse sentido os alunos desenvolvem a capacidade de monitorar seus próprios processos de aprendizagem e fazerem ajustes com base em suas reflexões, por exemplo. Assim, a avaliação como aprendizagem permite que os alunos se tornam mais independentes e ganhem confiança em suas próprias capacidades de controlar e direcionar o seu aprendizado.

Earl (2012) discute que os alunos também se demonstram como mais responsáveis ao engajarem com esse tipo de avaliação, pois refletem a respeito de seus pontos fortes e fracos, definem metas de aprendizagem e metas pessoais e consideram diferentes estratégias para melhorar.

Earl (2012) argumenta que integrar os três tipos de avaliação (avaliação da aprendizagem, avaliação para aprendizagem e avaliação como aprendizagem) pode levar a experiências de aprendizagem mais eficazes para os alunos, já que um ambiente de aprendizagem mais dinâmico, eficaz e flexível é promovido, indo além da mera medição do desempenho dos alunos, e fazendo os alunos alcançarem um desenvolvimento contínuo e autonomia no processo de aprendizagem.

3 ASPECTOS METODOLÓGICOS

Neste capítulo, relatamos o procedimento realizado nessa investigação para alcançar os objetivos da pesquisa. Também descrevemos os fundamentos de uma revisão bibliográfica sistemática e apresentamos um guia de oito passos para uma revisão bibliográfica sistemática, conforme Okoli (2015). Finalmente, explicitamos como os oito passos foram realizados no caso deste estudo e indicamos o referencial analítico dessa pesquisa: a Análise de Conteúdo de Bardin (2011), bem como os encaminhamentos analíticos assumidos nessa pesquisa.

3.1 REVISÃO BIBLIOGRÁFICA SISTEMÁTICA

De acordo com Fink (2005, p. 3), uma revisão bibliográfica sistemática é: “um método sistemático, explícito, abrangente e reproduzível para identificar, avaliar e sintetizar o corpo existente de trabalhos concluídos e registrados produzidos por pesquisadores e estudiosos”. Assim, uma revisão bibliográfica sistemática deve ser sistemática, adotando uma abordagem metodológica clara, que descreve explicitamente os procedimentos realizados. É necessário que a revisão seja abrangente ao incluir todo o material relevante e, similarmente, reproduzível por outros pesquisadores que queiram adotar a mesma abordagem ao examinar o tema.

Para uma revisão bibliográfica ser considerada sistemática, é essencial descrever de maneira clara e detalhada os procedimentos utilizados, permitindo que outros pesquisadores repliquem o processo e obter resultados semelhantes. Isso é particularmente importante em uma *Standalone Systematic Literature Review*, sendo uma revisão bibliográfica sistemática independente, onde o foco principal é a análise da literatura existente, sem a coleta ou análise de novos dados primários, que é o caso da revisão apresentada nesta tese.

Os procedimentos de uma revisão bibliográfica sistemática independente também podem ser comumente encontrados na seção “fundamentação teórica” de outros artigos, com a intenção de fornecer a base teórica e o contexto para um problema de pesquisa. Em artigos, essa seção é geralmente referida como “revisão de literatura” ou “levantamento bibliográfico”. Além disso, esses procedimentos também podem ser utilizados em uma revisão de literatura que

compõe um capítulo de uma tese, sendo frequentemente denominada “revisão de literatura da tese”.

Há várias razões para realizar uma revisão bibliográfica sistemática independente. Segundo Fink (2005), esses tipos de revisões podem ser realizados para descrever o conhecimento existente para a prática profissional, identificar projetos e técnicas eficazes, localizar especialistas em um campo específico e descobrir fontes não publicadas. Okoli (2015) destaca que, além de seu escopo e rigor metodológico distintos, essas revisões podem resumir resultados ou evidências existentes, apontar lacunas na pesquisa atual e oferecer uma base para futuras investigações. Petticrew e Roberts (2006) afirmam que essas revisões são valiosas porque podem influenciar ou aconselhar políticas e apoiar práticas profissionais.

Okoli (2015) vê o termo “sistemática” em “revisão bibliográfica sistemática” como um adjetivo qualitativo. Segundo ele, um adjetivo qualitativo descreve a natureza de algo de maneira que pode ser qualificada por “mais” ou “menos”. Portanto, uma revisão pode ser considerada mais sistemática, menos sistemática ou extremamente sistemática. Essa perspectiva é importante, pois ao tratar a palavra “sistemática” como um adjetivo classificatório, criamos uma distinção entre tipos de revisões bibliográficas. Assim, uma revisão bibliográfica sistemática é aquela que tem a intenção explícita de ser conduzida de maneira sistemática (Okoli, 2015).

Hart (1998) também enfatiza a importância da criticidade na revisão, argumentando que uma revisão bibliográfica sistemática não se limita a uma mera compilação, coleção ou resumo de estudos existentes. Em vez disso, a revisão deve contribuir de duas maneiras: sintetizando o material disponível e oferecendo uma análise crítica da teoria existente. Segundo Okoli (2015), o rigor acadêmico é alcançado apenas quando ambas as funções são devidamente cumpridas.

Okoli (2015) ressalta a importância das revisões bibliográficas independentes para a comunidade científica, pois elas economizam tempo e esforço dos pesquisadores que buscam uma síntese abrangente da literatura existente. Portanto, a revisão bibliográfica sistemática deve garantir a confiança dos leitores. Vom Brocke *et al.* (2009) argumentam que o processo de inclusão e exclusão de fontes precisa ser conduzido de forma transparente para assegurar a credibilidade da revisão. Dessa forma, outros pesquisadores podem avaliar a exaustividade da revisão e utilizar seus resultados em suas próprias pesquisas. Além disso, Okoli (2015) e Hart

(1998) discutem que, ao realizar uma revisão, o pesquisador deve ser claro sobre a escolha do tema, as razões para essa escolha e como a revisão fundamenta futuros estudos.

Okoli (2015) fornece um guia (Quadro 1) para a realização de uma revisão bibliográfica sistemática. Este guia descreve em detalhes oito etapas que garantem uma revisão rigorosa, capaz de resumir e discutir de maneira abrangente a literatura existente.

Quadro 1 - Um guia de oito passos para uma Revisão Bibliográfica Sistemática

1) Identificar o objetivo
A primeira etapa de qualquer revisão requer que os revisores identifiquem claramente a finalidade da revisão e as metas pretendidas. Isso é necessário para que a revisão seja transparente para os leitores.
2) Elaborar o protocolo e instruir a equipe
Para qualquer revisão que empregue mais de um revisor, os revisores precisam ser claros e estarem de acordo com o procedimento que seguirão. Isso requer um protocolo escrito e detalhado, bem como uma instrução para que todos os revisores possuam consistência em como eles executarão a revisão.
3) Aplicar um filtro prático
Essa etapa também é denominada de triagem para inclusão. Esta etapa requer que os revisores sejam transparentes sobre quais estudos eles consideraram para revisão e quais eles excluíram. Para os estudos excluídos, os revisores devem apresentar suas razões práticas para não os considerar. Os revisores também devem justificar como a revisão continua abrangente, mesmo com as exclusões, considerando os critérios práticos de exclusão.
4) Buscar literatura
Os revisores precisam ser transparentes e claros ao descreverem os detalhes de busca por literatura e precisam explicar e justificar como eles asseguraram a abrangência da pesquisa.
5) Extrair dados
Após os revisores terem identificado todos os estudos que devem ser incluídos na revisão, é necessário extrair sistematicamente as informações aplicáveis de cada estudo.
6) Avaliar a qualidade
Essa etapa também é denotada de triagem para exclusão. Nesse momento os revisores precisam explicar os critérios que foram utilizados para excluir documentos por qualidade insuficiente. Os pesquisadores devem classificar todos os trabalhos incluídos, de acordo com as metodologias de pesquisa ou outros critérios de sua escolha.
7) Sintetizar os estudos
Essa etapa também é conhecida como análise, envolve combinar os fatos extraídos dos estudos usando técnicas apropriadas, sejam quantitativas, qualitativas ou ambas.
8) Escrever a Revisão
Além dos princípios padrão a serem seguidos na escrita de trabalhos de pesquisa, o processo de uma revisão bibliográfica sistemática precisa ser relatado em

detalhes suficientes para que outros pesquisadores possam reproduzir de forma independente os resultados da revisão.

Fonte: extraído e adaptado de Okoli (2015, tradução nossa)

3.2 CONTEXTO DA PESQUISA

O presente trabalho visou realizar uma revisão bibliográfica sistemática de artigos envolvendo Avaliação na área de Ensino de Química conforme as oito etapas previamente descritas por Okoli (2015).

Na revisão bibliográfica sistemática em questão, a etapa 1 envolveu a definição do problema de pesquisa e as justificativas para a realização da revisão, apresentadas na Introdução. A etapa 2 consistiu na elaboração do protocolo da revisão, incluindo o cronograma das atividades de pesquisa (Quadro 2). Nesta etapa, também foram selecionados os referenciais metodológicos e analíticos a serem utilizados.

Quadro 2 - Cronograma das atividades de pesquisa

Descrição da Atividade	Período da realização
Leitura de Referenciais Teóricos sobre Avaliação, Análise de Conteúdo (Bardin, 2011) e os Referenciais Metodológicos (Okoli, 2015).	11/2021 – 01/2022
Definição do modelo de inventário a ser utilizado, discussões com a orientadora e um pesquisador internacional a respeito do que seria incluído e excluído dos inventários.	02/2022 – 03/2022
Aplicação do filtro prático.	04/2022
Busca na literatura.	05/2022 – 07/2022
Extração de dados dos artigos selecionados e avaliação da qualidade dos artigos.	08/2022 – 09/2022
Preenchimento, codificação e organização dos inventários.	09/2022 – 11/2022
Organização dos artigos em categorias conforme suas características gerais, seus objetivos e focos de avaliação.	12/2022 – 04/2023
Identificação das lacunas e tendências de pesquisa dos artigos analisados.	04/2023 – 05/2023
Discussões com a orientadora a respeito das lacunas e tendências de pesquisa. Formulação de hipóteses e definição dos caminhos a seguir para a escrita dos resultados.	06/2023
Primeira escrita dos resultados da Tese.	07/2023 – 01/2024
Discussões com a orientadora a respeito dos resultados da revisão e elaboração da entrevista com os principais pesquisadores em Avaliação na área de Ensino de Química.	02/2024 – 03/2024

Continuação da escrita da Tese	04/2024 – 05/2024
Entrevistas com os pesquisadores mais representativos em Avaliação na área de Ensino de Química.	07/2024 – 08/2024
Análise das entrevistas.	08/2024
Finalização da escrita da Tese.	08/2024

Fonte: Os autores

Destacamos que o cronograma das atividades de pesquisa funcionou como um guia inicial para a realização da revisão. No entanto, algumas atividades foram modificadas ou realizadas em períodos diferentes devido a imprevistos, ou ajustes. Portanto, o cronograma não deve ser visto como um relato preciso de como a pesquisa foi conduzida, mas sim como o protocolo inicial elaborado no início do processo. Incluímos o cronograma na tese para os leitores poderem acompanhar as alterações feitas na revisão ao longo do tempo. Os procedimentos finais realizados serão detalhados a seguir.

A etapa 3 consistiu na aplicação dos filtros e na definição dos critérios de exclusão e a etapa 4 consistiu na busca pela literatura. Para essa revisão, realizaram-se buscas em duas bases de dados: *Education Resources Information Center (ERIC)*¹⁷ e *European Reference Index for the Humanities (ERIH Plus)*¹⁸. Para todas as bases, buscamos a expressão “ASSESSMENT” e “CHEMISTRY EDUCATION” e selecionamos os seguintes filtros: artigos e artigos de revisão¹⁹; artigos de periódicos revisados por pares; e artigos publicados em 20 anos de busca (2002-2021)²⁰.

Essa primeira busca gerou um total de 252 resultados. Do total, 197 resultados foram do ERIC; e 55 resultados foram do ERIH Plus. Entretanto, houve um número considerável de artigos duplicados entre as bases de dados. Além disso, identificamos artigos que não poderiam prosseguir para as próximas etapas, por não estarem alinhados aos nossos filtros iniciais descritos nas etapas 3 e 4. Isso resultou na exclusão de 39 artigos, reduzindo o total de artigos para 213, conforme o Quadro 3.

¹⁷ <https://ERIC.ed.gov>

¹⁸ <https://kanalregister.hkdir.no/publiseringsskanaler/ERIHplus/>

¹⁹ Termo utilizado no Web of Science para buscar artigos que resumem o estado atual de entendimento de um tópico.

²⁰ Esse filtro não estava disponível no ERIC, logo tivemos que filtrar esse critério manualmente.

Quadro 3 - Resultados da busca nas bases de dados

Base de dados	Nº de artigos antes das exclusões	Nº de artigos excluídos	Nº de artigos final
ERIC	197	1	196
ERIH Plus	55	38	17
Total	252	39	213

Fonte: os autores

A fim de tornar todas as etapas da revisão bibliográfica sistemática transparentes e reproduzíveis, conforme recomendado por Okoli (2015), apresentamos abaixo a relação dos artigos excluídos no movimento supracitado e os respectivos motivos.

Quadro 4 - Artigos excluídos nas etapas 3 e 4

Nº	Base	Motivo	Referência
1	ERIC	Não é artigo	HOLME, T. A. What role may intuitive concepts about chemical ideas play when students take timed tests? Journal of Chemical Education , v. 98, n. 10, p. 3059-3060, 2021.
2	ERIH Plus	Duplicado ERIH Plus & ERIC	ANWAR, Y. A. S.; MUTI'AH, M. A. Exploration of the scientific papers and self-assessment of students using the COVID-19 case on biochemistry course. Biochemistry and Molecular Biology Education , v. 49, n. 3, p. 326-332, 2021.
3	ERIH Plus	Duplicado ERIH Plus & ERIC	ATIEH, E. L.; YORK, D. M.; MUÑIZ, M. N. Beneath the surface: an investigation of general chemistry students' study skills to predict course outcomes. Journal of Chemical Education , v. 98, n. 2, p. 281-292, 2020.
4	ERIH Plus	Não é artigo	HACH, B. Scholarships for High School Teachers. Journal of Chemical Education , v. 84, n. 12, p. 1896, 2007.
5	ERIH Plus	Duplicado ERIH Plus & ERIC	LAZENBY, K.; BALABANOFF, M. E.; BECKER, N. M.; MOON, A.; BARBERA, J. From Ideas to Items: A Primer on the Development of Ordered Multiple-Choice Items for Investigating the Progression of Learning in Higher Education STEM. Journal of Chemical Education , v. 98, n. 3, p. 714-729, 2021.
6	ERIH Plus	Não é artigo	PIENTA, N. J. This Is NOT Acceptable. Journal of Chemical Education , v. 91, n. 6, p. 771-772, 2014.
7	ERIH Plus	Duplicado ERIH Plus & ERIC	ELMGREN, M.; HO, F.; ÅKESSON, E.; SCHMID, S.; TOWNS, M. Comparison and evaluation of learning outcomes from an international perspective: Development of a best-practice process. Journal of Chemical Education , v. 92, n. 3, p. 427-432, 2015.

8	ERIH Plus	Duplicado ERIH Plus & ERIC	CHEN, K.; CHEN, Y.; LING, Y.; LIN, J. The individual experience of online chemistry teacher education in China: Coping with COVID-19 pandemic. Journal of Chemical Education , v. 97, n. 9, p. 3265-3270, 2020.
9	ERIH Plus	Duplicado ERIH Plus & ERIC	MURCIANO-CALLES, J. Use of Kahoot for assessment in chemistry education: a comparative study. Journal of Chemical Education , v. 97, n. 11, p. 4209-4213, 2020.
10	ERIH Plus	Duplicado ERIH Plus & ERIC	NEILES, K. Y.; TODD, I.; BUNCE, D. M. Establishing the validity of using network analysis software for measuring students' mental storage of chemistry concepts. Journal of Chemical Education , v. 93, n. 5, p. 821-831, 2016.
11	ERIH Plus	Duplicado ERIH Plus & ERIC	SEGEDINAC, M. D.; RODIC, D. D.; RONCEVIC, T. N.; HORVAT, S.; ADAMOV, J. The development of research in the field of chemistry education at the University of Novi Sad since the breakup of the Socialist Federal Republic of Yugoslavia. CEPS Journal , v. 10, n. 1, p. 103-124, 2020.
12	ERIH Plus	Duplicado ERIH Plus & ERIC	PRINS, G. T.; BULTE, A. M.; VAN DRIEL, J. H.; PILOT, A. Selection of authentic modelling practices as contexts for chemistry education. International Journal of Science Education , v. 30, n. 14, p. 1867-1890, 2008.
13	ERIH Plus	Duplicado ERIH Plus & ERIC	FOX, L. J.; ROEHRIG, G. H. Nationwide survey of the undergraduate physical chemistry course. Journal of Chemical Education , v. 92, n. 9, p. 1456-1465, 2015.
14	ERIH Plus	Duplicado ERIH Plus & ERIC	SCHWARTZ, A. T. Contextualized chemistry education: The AmERICan experience. International Journal of Science Education , v. 28, n. 9, p. 977-998, 2006.
15	ERIH Plus	Duplicado ERIH Plus & ERIC	RAJE, S.; STITZEL, S. Strategies for effective assessments while ensuring academic integrity in general chemistry courses during COVID-19. Journal of Chemical Education , v. 97, n. 9, p. 3436-3440, 2020.
16	ERIH Plus	Duplicado ERIH Plus & ERIC	LEWIS, S. E. Examining evidence for external and consequential validity of the first term general chemistry exam from the ACS Examinations Institute. Journal of Chemical Education , v. 91, n. 6, p. 793-799, 2014.
17	ERIH Plus	Duplicado ERIH Plus & ERIC	MACK, M. R.; HENSEN, C.; BARBERA, J. Metrics and methods used to compare student performance data in chemistry education research articles. Journal of Chemical Education , v. 96, n. 3, p. 401-413, 2019.
18	ERIH Plus	Duplicado ERIH Plus & ERIC	GRAULICH, N. Intuitive judgments govern students' answering patterns in multiple-choice exercises in organic chemistry. Journal of Chemical Education , v. 92, n. 2, p. 205-211, 2015.
19	ERIH Plus	Duplicado ERIH Plus & ERIC	HARSHMAN, J.; YEZIERSKI, E. Test-retest reliability of the adaptive chemistry assessment survey for teachers: Measurement error and alternatives to correlation. Journal of Chemical Education , v. 93, n. 2, p. 239-247, 2016.

20	ERIH Plus	Duplicado ERIH Plus & ERIC	TALANQUER, V. Chemical rationales: another triplet for chemical thinking. International Journal of Science Education , v. 40, n. 15, p. 1874-1890, 2018.
21	ERIH Plus	Duplicado ERIH Plus & ERIC	DAVENPORT, J. L.; RAFFERTY, A. N.; YARON, D. J. Whether and how authentic contexts using a virtual chemistry lab support learning. Journal of Chemical Education , v. 95, n. 8, p. 1250-1259, 2018.
22	ERIH Plus	Duplicado ERIH Plus & ERIC	SZTEINBERG, G. et al. Collaborative professional development in chemistry education research: bridging the gap between research and practice. Journal of Chemical Education , v. 91, n. 9, p. 1401-1408, 2014.
23	ERIH Plus	Duplicado ERIH Plus & ERIC	O'DWYER, A.; CHILDS, P. Organic chemistry in action! Developing an intervention program for introductory organic chemistry to improve learners' understanding, interest, and attitudes. Journal of Chemical Education , v. 91, n. 7, p. 987-993, 2014.
24	ERIH Plus	Duplicado ERIH Plus & ERIC	HOLME, T. A.; LUXFORD, C. J.; BRANDRIET, A. Defining conceptual understanding in general chemistry. Journal of Chemical Education , v. 92, n. 9, p. 1477-1483, 2015.
25	ERIH Plus	Duplicado ERIH Plus & ERIC	HANSON, R. Employing Microscience Equipment to Promote Chemistry Education Through Constructivist Hands-and Minds-on Activities. Interchange , v. 52, n. 4, p. 521-544, 2021.
26	ERIH Plus	Duplicado ERIH Plus & ERIC	HOLME, T.; MURPHY, K. Assessing conceptual and algorithmic knowledge in general chemistry with ACS exams. Journal of Chemical Education , v. 88, n. 9, p. 1217-1222, 2011.
27	ERIH Plus	Duplicado ERIH Plus & ERIC	RAVI, M.; PUENTE-URBINA, A.; VAN BOKHOVEN, J. A. Identifying Opportunities to Promote Systems Thinking in Catalysis Education. Journal of Chemical Education , v. 98, n. 5, p. 1583-1593, 2021.
28	ERIH Plus	Duplicado ERIH Plus & ERIC	ARMSTRONG, L. B. et al. Developing a Green Chemistry Focused General Chemistry Laboratory Curriculum: What Do Students Understand and Value about Green Chemistry?. Journal of Chemical Education , v. 96, n. 11, p. 2410-2419, 2019.
29	ERIH Plus	Duplicado ERIH Plus & ERIC	STOWE, R. L. et al. You are what you assess: The case for emphasizing chemistry on chemistry assessments. Journal of Chemical Education , v. 98, n. 8, p. 2490-2495, 2021.
30	ERIH Plus	Duplicado ERIH Plus & ERIC	JAGODZINSKI, P.; WOLSKI, R. Assessment of application technology of natural user interfaces in the creation of a virtual chemical laboratory. Journal of Science Education and Technology , v. 24, n. 1, p. 16-28, 2015.
31	ERIH Plus	Duplicado ERIH Plus & ERIC	SCHULTZ, M. et al. Evaluation of diagnostic tools that tertiary teachers can apply to profile their students' conceptions. International Journal of Science Education , v. 39, n. 5, p. 565-586, 2017.

32	ERIH Plus	Duplicado ERIH Plus & ERIC	MA, J. Getting a Bigger Picture in Less Time: Viewing Curriculum Reform in a Chinese Graduate Chemistry Program through the Lens of an Organic Structure Analysis Course. Journal of Chemical Education , v. 88, n. 12, p. 1639-1643, 2011.
33	ERIH Plus	Duplicado ERIH Plus & ERIC	COLLINS, J. S.; OLESIK, S. V. The important role of chemistry department chairs and recommendations for actions they can enact to advance black student success. Journal of Chemical Education , v. 98, n. 7, p. 2209-2220, 2021.
34	ERIH Plus	Duplicado ERIH Plus & ERIC	RUSMINI, R.; SUYONO, S.; AGUSTINI, R. Analysis of science process skills of chemical education students through self project based learning (SjBL) in the pandemic COVID 19 era. JOTSE , v. 11, n. 2, p. 371-387, 2021.
35	ERIH Plus	Duplicado ERIH Plus & ERIC	SCHENK, L.; TAHER, I. A.; ÖBERG, M. Identifying the scope of safety issues and challenges to safety management in Swedish middle school and high school chemistry education. Journal of Chemical Education , v. 95, n. 7, p. 1132-1139, 2018.
36	ERIH Plus	Duplicado ERIH Plus & ERIC	EATON, A. C.; DELANEY, S.; SCHULTZ, M. Situating sustainable development within secondary chemistry education via systems thinking: a depth study approach. Journal of Chemical Education , v. 96, n. 12, p. 2968-2974, 2019.
37	ERIH Plus	Duplicado ERIH Plus & ERIC	SCOTT, D.; FIRTH, D. Using control charts early in the quantitative analysis laboratory curriculum. Journal of Chemical Education , v. 96, n. 5, p. 1037-1041, 2019.
38	ERIH Plus	Não é artigo	HOLME, T. Assessment and quality control in chemistry education. Journal of Chemical Education , v. 80, n. 6, p. 594, 2003.
39	ERIH Plus	Não é artigo	BRETZ, S. L. Implementing the professional development standards: an innovative MS degree for high school chemistry teachers. Journal of Chemical Education , v. 79, n. 11, 2002.

Fonte: os autores

Na etapa 5, a fim de extrair os dados relevantes para análise dos artigos, optamos por utilizar inventários. Um inventário foi preenchido para cada artigo a fim de possibilitar o mapeamento das características dos artigos envolvendo Avaliação na área de Ensino de Química, tal como seguido em outros estudos similares (Costa; Broietti, 2022; Costa *et al.*, 2021; Costa; Broietti, 2021a; Costa; Broietti, 2021b; Costa; Obara; Broietti, 2020a; Costa; Obara; Broietti, 2020b). O modelo do inventário utilizado é apresentado no Quadro 5.

Quadro 5 - Modelo de Inventário utilizado

Código	
Referência APA	
Instituição dos autores	
País	
Abstract	
Objetivo(s)/Pergunta(s) de pesquisa	
Nível de Ensino	
Palavras-chave	
Citações	

Fonte: os autores

Os inventários foram uma ferramenta chave para permitir a análise dos artigos. Como o volume do *corpus* era elevado, comparações de informações específicas entre artigos e a própria categorização dos mesmos não seria possível com os artigos na íntegra, devido à própria extensão do *corpus* inicial. Assim, os inventários serviram para reduzir cada artigo em somente suas informações-chave, ou necessárias, para identificarmos suas principais características e realizarmos as categorizações.

Após preenchermos os inventários dos artigos, realizamos uma primeira leitura deles para avaliarmos a qualidade. Nesse processo, 8 artigos foram excluídos por não possuírem um enfoque suficiente em Avaliação. Dessa forma, o *corpus* de nossa pesquisa foi reduzido de 213 artigos para 205 artigos, constituindo o *corpus* final da tese. Ressaltamos que nenhuma outra exclusão foi realizada no *corpus*. Isso concluiu a etapa 6. A fim de tornarmos todas as exclusões transparentes e demonstrarmos a abrangência do *corpus*, apresentamos as últimas exclusões no Quadro 6:

Quadro 6 - Artigos excluídos na etapa 6

Motivo da exclusão	Referência (Código do artigo na primeira codificação)
Não possui enfoque suficiente em avaliação ²¹ . Discute e caracteriza os principais tipos de raciocínios construídos e aplicados por especialistas ao analisar sistemas e processos químicos. Não avalia algo ²² . Apenas comenta no abstract que o	TALANQUER, V. Chemical rationales: another triplet for chemical thinking. <i>International Journal of Science Education</i> , v. 40, n. 15, p. 1874-1890, 2018. (Antigo A036)

²¹ Fizemos esse julgamento baseado em alguns fatores. Se os artigos continham o termo “avaliação”, “avalia”, “avaliar”, etc. em seus objetivos ou perguntas de pesquisa. Também, se a avaliação era um dos aspectos centrais do trabalho, que foi verificado a partir da leitura, interpretação e contagem e análise do uso dos termos termo “avaliação”, “avalia”, “avaliar” no corpo do texto.

²² Notamos que o artigo paira apenas na discussão e caracterização e não adentra na avaliação.

artigo é importante para subsidiar a elaboração de avaliações que fomentem essas formas de pensar.	
Não possui enfoque suficiente em avaliação. Discute e elabora uma articulação/definição consensual de compreensão conceitual definida pelos formadores. Não avalia algo ²³ . Apenas comenta no abstract que o artigo é importante, pois a compreensão conceitual é medida em avaliações.	HOLME, T. A.; LUXFORD, C. J.; BRANDRIET, A. Defining conceptual understanding in general chemistry. Journal of Chemical Education , v. 92, n. 9, p. 1477-1483, 2015. (Antigo A049)
Não possui enfoque suficiente em avaliação. O artigo apenas discute que os resultados da pesquisa fornecem informações sobre estratégias de avaliação para tarefas semelhantes.	BECKER, N. M.; RUPP, C. A.; BRANDRIET, A. Engaging students in analyzing and interpreting data to construct mathematical models: An analysis of students' reasoning in a method of initial rates task. Chemistry Education Research and Practice , v. 18, n. 4, p. 798-810, 2017. (Antigo A072)
Não possui enfoque suficiente em avaliação. O artigo apenas discute que os resultados fornecem implicações para elaborar avaliações no abstract.	BRESLYN, W. PAA black holes, Khan, and Quora: mapping understanding of isotopes through existing data sources. Chemistry Education Research and Practice , v. 21, n. 1, p. 412-425, 2020. (Antigo A128)
Não possui enfoque suficiente em avaliação. O artigo apenas discute o Ensino de Química baseado em contexto nos Estados Unidos (como um currículo).	SCHWARTZ, A. T. Contextualized chemistry education: The AmERICan experience. International Journal of Science Education , v. 28, n. 9, p. 977-998, 2006. (Antigo A045)
Não possui enfoque suficiente em avaliação. Apenas comenta brevemente que é necessário compreender o pensamento química dos alunos para desenvolver avaliações autênticas.	SEVIAN, H.; TALANQUER, V. Rethinking chemistry: A learning progression on chemical thinking. Chemistry Education Research and Practice , v. 15, n. 1, p. 10-23, 2014. (Antigo A147)
Não possui enfoque suficiente em avaliação. O artigo analisa a segurança de laboratórios de escolas e comenta que há uma falta de avaliações de risco, escritas nos laboratórios, que deveriam garantir a segurança dos alunos. Dentre os outros aspectos, são mencionados: instruções de	SCHENK, L.; TAHER, I. A.; ÖBERG, M. Identifying the scope of safety issues and challenges to safety management in Swedish middle school and high school chemistry education. Journal of Chemical

²³ Notamos que o artigo paira apenas na discussão e elaboração e não adentra na avaliação.

manuseio seguro mal documentados; armazenamento de produtos químicos inadequados, funcionalidade e tamanho insuficientes das salas de aula, grupos de alunos muito grandes e procedimentos de primeiros socorros ausentes.	Education , v. 95, n. 7, p. 1132-1139, 2018. (Antigo A100)
Não possui enfoque suficiente em avaliação. O artigo descreve uma colaboração entre professores e pesquisadores educacionais. O artigo apenas comenta brevemente, ao fim do trabalho, que a análise revelou alguns resultados, como por exemplo, as avaliações dos professores mudaram para uma prática mais baseada em pesquisa.	SZTEINBERG, G. et al. Collaborative professional development in chemistry education research: bridging the gap between research and practice. Journal of Chemical Education , v. 91, n. 9, p. 1401-1408, 2014. (Antigo A042)

Fonte: os autores

Com a saída desses 8 artigos, o *corpus* teve que ser recodificado de A001-A213 para A001-A205. Isso foi realizado por meio da substituição do código dos últimos artigos (A206, A207, A208, A209, A210, A211, A212, A213), na primeira codificação, para os códigos dos artigos recém-excluídos (A036, A049, A072, A0128, A045, A147, A100, A042).

Assim, o artigo A206 passou a ser chamado de A036, por exemplo, substituindo o lugar de um artigo eliminado. Todas as substituições de códigos entre o primeiro e segundo movimento de codificação podem ser visualizadas no Quadro 7.

Quadro 7 - Recodificação por razão das exclusões

Código do artigo na primeira codificação	Código do artigo na segunda codificação
A206	A036
A207	A049
A208	A072
A209	A128
A210	A045
A211	A147
A212	A100
A213	A042

Fonte: os autores

Diante do *corpus* da pesquisa, os 205 artigos, para realizar a etapa 7, Okoli (2015) recomenda o uso de técnicas apropriadas, sejam essas qualitativas ou quantitativas. Para essa tese, foi utilizada a Análise de Conteúdo proposta por Bardin (2011). Na subseção a seguir, apresentamos alguns pressupostos da Análise de

Conteúdo em maiores detalhes e os motivos pelos quais essa técnica foi selecionada para esta revisão.

3.3 ANÁLISE DE CONTEÚDO

Para Bardin (2011), existem correspondências entre as estruturas semânticas ou linguísticas do texto e as estruturas psicológicas ou sociológicas do mesmo. Assim, inferências podem ser feitas em relação ao texto. Essa ideia está presente na própria definição da análise de conteúdo de Bardin (2011), concebida como:

Um conjunto de técnicas de análise das comunicações visando obter, por procedimentos sistemáticos e objetivos de descrição do conteúdo das mensagens, indicadores (quantitativos ou não) que permitam a inferência de conhecimentos relativos às condições de produção/recepção (variáveis inferidas) destas mensagens (p. 48, grifo do autor).

Nesse sentido, Bardin (2011) explicita o que se entende pelo termo “condições de produção” e as diversas variáveis que isso pode contemplar:

O termo *condições de produção* é suficientemente vago para permitir possibilidades de inferência muito variadas: variáveis psicológicas do indivíduo emissor, variáveis sociológicas e culturais, variáveis relativas à situação de comunicação ou do contexto de produção da mensagem (p. 46, grifo do autor).

Logo, na Análise de Conteúdo, por meio do tratamento das mensagens que o analista manipula, o conhecimento a respeito do emissor da mensagem ou do meio é inferido. A Análise de Conteúdo é estruturada em três etapas: 1) *A pré-análise*; 2) *A exploração do material*; 3) *O tratamento dos resultados, a inferência e a interpretação*.

A *pré-análise* é uma etapa de organização, na qual o objetivo é sistematizar as ideias iniciais e tornar o material operacional. A *pré-análise* inclui a leitura flutuante, ou seja, o primeiro contato com o material, em que se começa a conhecer o texto; a seleção dos documentos; a formulação das hipóteses e a definição dos objetivos; a referenciação dos índices/elaboração de indicadores, que envolve determinar indicadores por meio de recortes no material de análise e a preparação do material (Bardin, 2011). Neste estudo, a *pré-análise* consistiu no primeiro contato com

os artigos, momento em que foi realizada a primeira leitura, e a extração das informações necessárias de cada artigo para preencher os inventários.

A segunda etapa, a *exploração do material*, é caracterizada pela administração sistemática das decisões tomadas anteriormente. Esta fase, considerada longa e fastidiosa, pela autora, consiste em operações de codificação e enumeração em função de regras previamente formuladas. Nesta fase é necessário um estudo, guiado pelas hipóteses e referenciais teóricos do estudo. A *exploração do material* inclui a codificação, a classificação e a categorização (Bardin, 2011). A codificação corresponde a uma transformação como: recorte, agregação e enumeração, para alcançar uma representação do conteúdo ou da sua expressão.

Ao finalizar a codificação, inicia-se a categorização, que conforme Bardin:

[...] é uma operação de classificação de elementos constitutivos de um conjunto, por diferenciação e, seguidamente, por reagrupamento segundo o género (analogia), com os critérios previamente definidos. As categorias, são rubricas ou classes, as quais reúnem um grupo de elementos (unidades de registo, no caso da análise de conteúdo) sob um título genérico, agrupamento esse efectuado em razão dos caracteres comuns destes elementos (Bardin, 1977, p. 117).

Neste estudo, a *exploração do material* consistiu na codificação dos artigos de A001-A205; o agrupamento dos artigos conforme suas características similares; e a categorização dos artigos conforme semelhanças no que havia sido avaliado. Apresentamos os resultados desses processos na seção de Resultados e Discussão dessa tese.

Finalmente, na terceira etapa, *O tratamento dos resultados, a inferência e a interpretação*, o analista deve propor inferências e adiantar interpretações a propósito dos objetivos previstos (Bardin, 2011). Nesta tese, esta etapa consistiu em apresentar os resultados das categorizações e similaridades encontradas entre os artigos quanto às suas características e o que havia sido avaliado, além da discussão dos resultados.

Logo, a etapa 7 envolveu as três fases da Análise de Conteúdo de Bardin (2011). A última etapa de Okoli (2015), a etapa 8, consistiu na escrita da presente tese e apresentação de suficientes detalhes para assegurar a replicabilidade independente do estudo.

3.4 ENCAMINHAMENTOS METODOLÓGICOS E ANALÍTICOS DAS ENTREVISTAS

Resgatamos o terceiro objetivo de pesquisa: identificar e analisar as perspectivas dos principais autores internacionais em Avaliação na área de Ensino de Química sobre as características, tendências e as lacunas das pesquisas. Enviamos o convite para a participação nas entrevistas para nove pesquisadores, após identificarmos os pesquisadores mais proeminentes no tema investigado. Entretanto, obtivemos retorno de apenas quatro dos nove pesquisadores. Logo, conduzimos entrevistas com esses quatro pesquisadores, codificados como P1 a P4 com base na ordem alfabética de seus nomes.

A coleta de dados foi realizada por meio de dois métodos principais: respostas escritas e entrevistas síncronas. Dois participantes optaram por oferecer suas percepções por meio de respostas escritas estruturadas, enquanto outros dois preferiram realizar entrevistas síncronas, que foram realizadas por plataformas de videoconferência. Todos os participantes assinaram um termo de consentimento para participação na pesquisa.

Apresentamos, no Quadro 8, as informações a respeito dos participantes das entrevistas e como se deu a tomada de dados.

Quadro 8 - Participantes das entrevistas e coleta de dados

Pesquisadores entrevistados	Universidade	Coleta de dados
P1	Miami University	Entrevista síncrona com duração de trinta e três minutos, no dia 09/07/2024.
P2	Clemson University	Entrevista síncrona com duração de trinta e nove minutos, no dia 15/08/2024.
P3	University of South Florida	Respondeu ao questionário de forma escrita.
P4	Technion-Israel Institute of Technology	Respondeu ao questionário de forma escrita.

Fonte: os autores

As entrevistas síncronas, foram gravadas usando tecnologia de áudio e vídeo e posteriormente transcritas na íntegra para possibilitar uma análise detalhada.

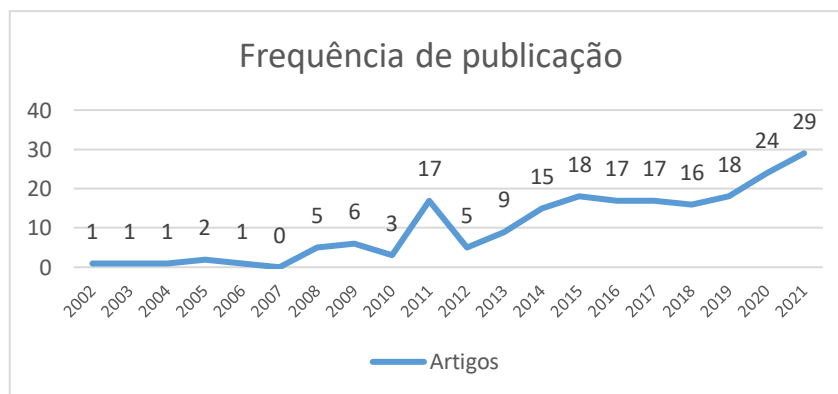
As questões propostas aos pesquisadores encontram-se no Quadro 9. Destacamos que estas foram apresentadas ao grupo de pesquisa EDUCIM, a qual

pertencemos e foram discutidas/validadas, trazendo maior confiabilidade as nossas perguntas.

Quadro 9 – Perguntas da Entrevista²⁴

Entrevista

1. O que você entende por “avaliação”?
2. Como você vê a Avaliação como um campo de pesquisa no Ensino de Química?
3. O que tem sido frequentemente avaliado, talvez excessivamente, no Ensino de Química? Por que você acredita que esse é o caso?
4. O que não tem sido avaliado, o suficiente, no Ensino de Química? Por que você acredita que esse é o caso?
5. Nossos resultados de uma revisão bibliográfica sistemática sobre Avaliação na área de Ensino de Química, em duas décadas, mostram uma predominância na avaliação, no nível do Ensino Superior, em Ensino de Química. Por que você acredita que esse é o caso?
6. Nossos resultados de uma revisão bibliográfica sistemática sobre Avaliação na área de Ensino de Química, em duas décadas, mostram haver um interesse crescente em pesquisar Avaliação na área de Ensino de Química (como visto na figura abaixo). A que você atribui esse interesse crescente?



7. Nossos resultados de uma revisão bibliográfica sistemática sobre Avaliação na área de Ensino de Química, em duas décadas, resultaram em 10 categorias. Essas 10 categorias revelam o que foi avaliado nos 205 artigos investigados. Qual é sua opinião sobre o que foi avaliado em Ensino de Química, conforme visto na tabela abaixo?

²⁴ Apresentamos nesta tese as perguntas traduzidas para a Língua Portuguesa, porém estas foram realizadas em Língua Inglesa no momento das entrevistas, devido aos participantes serem do contexto internacional.

Categoria	Nº de artigos	Porcentagem
C01: Avalia habilidades	18	8,8%
C02: Avalia estratégias de avaliação ²⁵ , provas, exames e itens	17	8,3%
C03: Avalia propostas de ensino	39	19,0%
C04: Avalia um instrumento, ferramenta ou tecnologia ²⁶	36	17,5%
C05: Avalia compreensões, percepções e interesse	19	9,3%
C06: Avalia conhecimentos	20	9,8%
C07: Avalia o currículo	14	6,8%
C08: Avalia atitudes	10	4,9%
C09: Avalia outras pesquisas	3	1,5%
C10: Discute avaliação em contextos diferentes	29	14,1%
Total	205	100%

8. Na sua opinião, quais são os maiores obstáculos/dificuldades da avaliação em sala de aula hoje?

Fonte: os autores

Para aumentar a clareza da análise e tornar as respostas dos pesquisadores gerenciáveis, principalmente as respostas derivadas das entrevistas síncronas, as respostas de cada participante foram fragmentadas em blocos menores de texto. Essa abordagem nos permitiu isolar temas e ideias específicas de cada participante, facilitando a análise das diferentes percepções.

Para a análise dos dados das entrevistas, utilizamos a Análise de Conteúdo, seguindo a metodologia delineada por Bardin (2011). Escolhemos a Análise de Conteúdo, por ser uma técnica sistemática e objetiva que nos permitiu quantificar e analisar a presença de temas semelhantes manifestados pelos pesquisadores e/ou conceitos específicos em dados qualitativos. A Análise de Conteúdo é dividida em três etapas: *pré-análise*, *exploração do material*, *tratamento dos resultados*. Na sequência, fornecemos o passo a passo de como isso foi realizado conforme a Análise de Conteúdo.

²⁵ Traduzido do Inglês *assessments*. Vários artigos utilizam o termo *assessments*. *Assessment*, em inglês, possui um sentido de prova mais formal, as vezes aplicada em larga escala, contrastado com o termo *test*, que seria a forma mais casual de prova.

²⁶ Todos os instrumentos ou ferramentas que envolvem mídia, foram deslocados para a subcategoria de tecnologias.

Neste estudo, a *pré-análise* consistiu na transcrição das entrevistas para a organização e formatação para posterior análise. Também consistiu no primeiro contato com as respostas dos pesquisadores, no qual foi realizada a primeira leitura, e a familiarização com suas percepções gerais.

A *exploração do material* consistiu na: codificação dos pesquisadores de P1 a P4; a fragmentação em blocos manejáveis das respostas dos pesquisadores, e o agrupamento das respostas conforme suas características similares. Uma vez que alguns temas principais foram encontrados nas respostas dos pesquisadores, conduzimos uma análise detalhada para realizarmos interpretações. Isso incluiu examinar a frequência de certos temas, identificar divergências e convergências nas percepções dos pesquisadores e considerar as implicações dessas descobertas para práticas de Avaliação na área de Ensino de Química.

O *tratamento dos resultados* consistiu em apresentar os resultados dos agrupamentos e similaridades encontradas entre as respostas dos pesquisadores quanto às suas percepções a respeito da Avaliação na área de Ensino de Química e os resultados que obtivemos da nossa pesquisa. Apresentamos os resultados desses processos na seção de Resultados e Discussão dessa tese.

Por meio desse encaminhamento metodológico, buscamos uma compreensão abrangente das diversas percepções expressas pelos principais pesquisadores sobre o tema em estudo.

4 RESULTADOS E DISCUSSÃO

Nesse capítulo, apresentamos os artigos analisados nesse estudo, assim como a análise das entrevistas com os principais autores internacionais em Avaliação, e tivemos em vista responder a cada uma das perguntas de pesquisa:

- I) Quais são as principais características, tendências e lacunas dos artigos internacionais sobre Avaliação na área de Ensino de Química?
- II) O que tem sido avaliado nos artigos internacionais sobre Avaliação na área de Ensino de Química?
- III) Quais são as perspectivas dos principais autores internacionais em Avaliação na área de Ensino de Química sobre as características, tendências e lacunas das pesquisas?

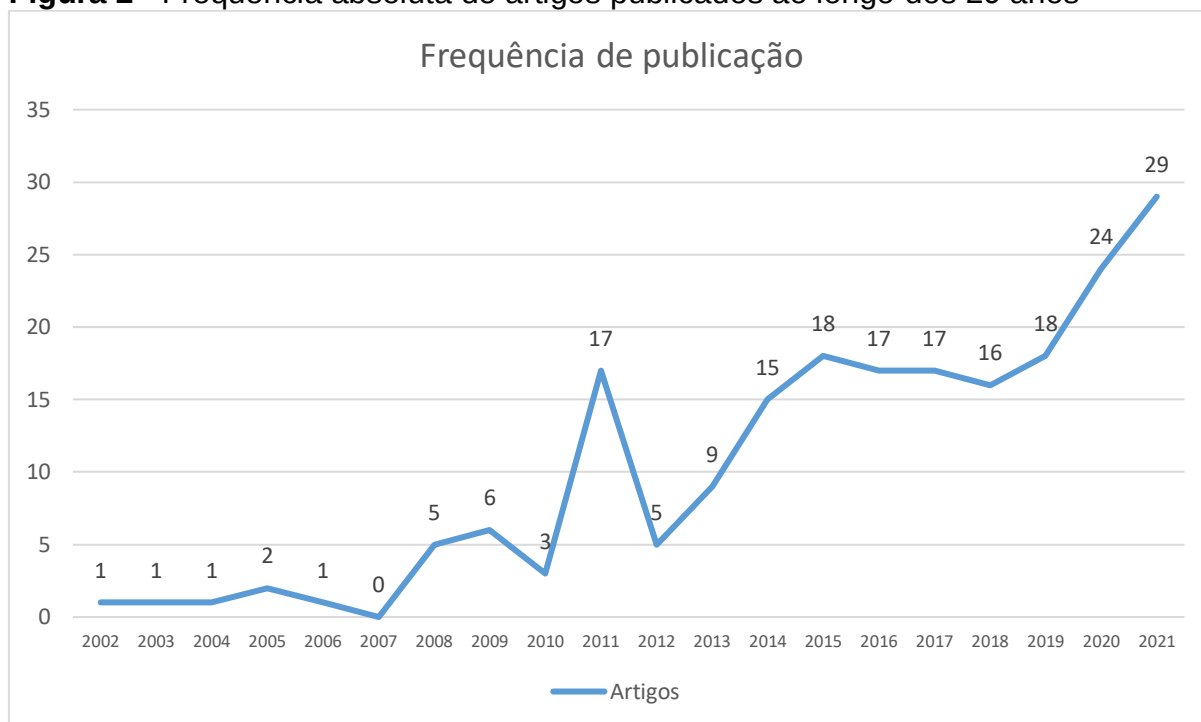
Apresentamos na Apêndice A, a codificação dos 205 artigos analisados na pesquisa. A primeira coluna corresponde ao código e a segunda coluna à referência bibliográfica do artigo. As discussões nesta seção foram realizadas utilizando os códigos dos artigos.

A seguir, apresentamos resultados e discussões de nossas análises, subdivididas em quatro seções, estas correspondentes a cada uma das questões de pesquisa propostas para este estudo.

4.1 QUAIS SÃO AS PRINCIPAIS CARACTERÍSTICAS DOS ARTIGOS INTERNACIONAIS SOBRE AVALIAÇÃO NA ÁREA DE ENSINO DE QUÍMICA?

Ao analisarmos as características das publicações, concentramos nossa atenção nos seguintes aspectos: o período de publicação; os autores; os periódicos; os países; os continentes; os níveis de ensino avaliados; e as instituições.

Em relação ao número de artigos publicados e o período de publicação, apresentamos a Figura 2.

Figura 2 - Frequência absoluta de artigos publicados ao longo dos 20 anos

Fonte: os autores

Ao analisar o gráfico da frequência de publicação, algumas observações podem ser feitas. A primeira é de que a maioria dos artigos sobre Avaliação na área de Ensino de Química foram publicadas na segunda metade do recorte temporal feito para esta revisão, de 2011-2021. A primeira década (2002-2011) contou com apenas 18% das publicações (37 artigos). Já a segunda década (2012-2021) contou com 82% das publicações (168 artigos). Logo, houve um aumento de interesse e disseminação de artigos desse tema na última década.

O aumento de artigos sobre Avaliação na área de Ensino de Química nos vinte anos selecionados como recorte (2002-2021) pode ser atribuído a um conjunto de fatores, estes que refletem a evolução das práticas e prioridades da comunidade educacional nesse período.

A mudança para práticas baseadas em evidências, visto principalmente na literatura internacional em Ensino de Química, pode ser um desses fatores. A ênfase em práticas de ensino e avaliações baseadas em evidências têm ganhado significativa força nas últimas duas décadas, refletindo uma mudança de paradigma nas abordagens educacionais (Hattie, 2009; Kuh *et al.*, 2005). Para Boud e Falchikov (2006), isso tem ocorrido devido à crescente demanda por transparência e responsabilidade das instituições de ensino, principalmente as de ensino superior.

Nesse sentido, os professores têm sido cada vez mais cobrados a demonstrar a qualidade ou eficácia de suas práticas de ensino e sua instrução por meio de resultados mensuráveis e dados concretos. Por esse motivo, a pesquisa em Ensino de Química pode ter respondido a essa crescente demanda com um incremento no número de artigos que avaliam intervenções em salas de aulas, abordagens, métodos, instruções, avaliações, etc. Essa hipótese é corroborada por um de nossos resultados a respeito do que a área tem mais valorizado nos artigos sobre Avaliação na área de Ensino de Química. Isso, pois a categoria mais representativa do nosso *corpus* foi a C03: Avalia propostas de ensino, representando 19% do *corpus*.

Outro fator que pode ter contribuído para o aumento da quantidade de artigos sobre Avaliação na área de Ensino de Química é a introdução do conceito de avaliações formativas por Black e William (1998) e o subsequente interesse da comunidade educacional em rever o papel da avaliação no contexto educacional. As avaliações formativas, em sua própria conceituação, permitem um acompanhamento mais próximo do progresso na aprendizagem dos estudantes e possibilitam ajustes imediatos nas práticas de ensino adotadas pelo professor. Além disso, os resultados dos alunos nas avaliações podem ser levados em consideração na tomada de decisões pedagógicas no futuro, permitindo que os professores ajustem suas abordagens para melhor alcançar os objetivos de aprendizagem e adequá-las para o contexto educacional em que eles estão inseridos (Morrison *et al.*, 2017).

Pensamos que o interesse em propor avaliações formativas pode ter contribuído para o aumento do interesse em publicações sobre Avaliação na área de Ensino de Química, já que esta se adequa melhor a ideia de uma aprendizagem centrada no aluno, uma mudança de paradigma na educação que não impactou apenas o Ensino de Química, mas toda a área de educação. A mudança da aprendizagem centrada no professor para a aprendizagem centrada no aluno tem sido uma transformação significativa nas práticas educacionais contemporâneas. Esse novo enfoque enfatiza a participação ativa dos alunos em seus próprios processos de aprendizagem, configurando-os como protagonistas no ambiente escolar. Acreditamos que esse movimento em direção a práticas de avaliação centradas no aluno tem contribuído para o rico corpo de pesquisa sobre Avaliação na área de Ensino de Química.

Nesse sentido, Boud e Falchikov (2006) discutem que as avaliações, saíram de um papel de instrumentos de medição de conhecimento e passaram a ser

vistas como oportunas para promover compreensões mais aprofundadas sobre os temas/conceitos estudados. Assim, as avaliações têm sido reformuladas para incluir a reflexão crítica, a conexão entre teoria e prática e a aplicação dos conceitos em situações ou contextos novos.

Os resultados de Broietti, Filho e Passos (2015) sobre avaliação na área de Ensino de Química em artigos nacionais revelam que essa é uma característica no contexto nacional também. A pesquisa supracitada analisou 17 artigos publicados entre 1986-2012 e verificou que mais de 90% dos artigos tinham sido publicados nos últimos nove anos do recorte temporal.

Em relação aos autores internacionais que mais publicaram artigos sobre Avaliação na área de Ensino de Química, destacamos o nome e a Instituição no Quadro 10.

Quadro 10 - Principais autores que publicaram em Avaliação na área de Ensino de Química

Nº de artigos	Autor	Instituição do autor
7	Ellen J. Yeziarski	Miami University
6	Scott E. Lewis	University of South Florida
6	Jennifer E. Lewis	University of South Florida
6	Melanie M. Cooper	Clemson University
5	Alison B. Flynn	University of Ottawa
4	Jack Barbera	Portland State University
4	Thomas Holme	University of Wisconsin
4	Marcy Towns	Purdue University
4	Yehudit Judy Dori	Technion-Israel Institute of Technology

Fonte: os autores

Optamos por fazer o corte dos principais autores em 4 artigos, pois havia um grande número de autores que publicaram apenas 3 artigos sobre Avaliação na área de Ensino de Química. Os autores apresentados no Quadro 10 destacaram-se pela sua produtividade e prolixidade, em comparação aos outros autores do *corpus*, no que diz respeito ao seu interesse em publicar pesquisas sobre Avaliação na área de Ensino de Química.

Ao analisarmos o Quadro 10, é possível notar que a maior parte dos principais autores em Avaliação na área de Ensino de Química dos artigos internacionais encontra-se nos Estados Unidos. 78% dos principais autores (sete autores) estão associados às instituições nos Estados Unidos e apenas 22% dos

principais autores (dois autores) encontram-se fora, respectivamente, no Canadá e Israel.

Além disso, nota-se uma considerável concentração de autores importantes em Avaliação na área de Ensino de Química no estado da Flórida, encontrado no litoral sudeste dos Estados Unidos. 33% dos principais autores (três autores) estão em instituições neste estado, na *Miami University* e na *University of South Florida*.

Ao localizarmos os principais autores, entramos em contato com todos os apresentados no Quadro 10, a fim de coletar suas perspectivas sobre as pesquisas na área de Avaliação na área de Ensino de Química.

Em relação aos periódicos nos quais os artigos foram publicados, apresentamos o Quadro 11.

Quadro 11 - Periódicos que os artigos foram publicados

Nº de Artigos	Periódico
136	Chemistry Education Research and Practice
32	Journal of Chemical Education
4	Journal of Science Education and Technology
4	International Journal of Science Education
1	Asia-Pacific Forum on Science Learning & Teaching Assessment Update Biochemistry and Molecular Biology Education CEPS Journal Computers & Education Educación química Education and Information Technologies Educational Action Research Educational Assessment, Evaluation and Accountability Electronic journal of e-Learning European Journal of Educational Research Higher education Higher Education Management and Policy Higher Education Policy Innovative Higher Education Interchange International Education Studies International Journal of Educational Development International Journal of STEM Education Issues in Educational Research JOTSE Journal of Education and Practice Journal of Educational Technology Journal of University Teaching & Learning Practice Learning and Instruction

	Pedagogical Research Research Papers in Education Science Education International Thinking Skills and Creativity
--	---

Fonte: os autores

Ao observar o Quadro 11, nota-se que a maioria dos artigos (66%) foram publicados na revista *Chemistry Education Research and Practice*, precedido pela *Journal of Chemical Education*, contando com 16% dos artigos. Logo, 82% de todas as pesquisas em Avaliação na área de Ensino de Química pertenceram a essas duas revistas, assumindo uma significativa representatividade dentro do *corpus*.

A *Chemistry Education Research and Practice*, conforme sua própria descrição em seu site, é um periódico para professores, pesquisadores e outros profissionais em todos os níveis da Educação em Química. O periódico é de livre acesso, sendo publicado gratuitamente, de forma eletrônica, quatro vezes ao ano. De acordo com o 2023 Journal Citation Reports (Clarivate Analytics, 2024) seu fator de impacto é 2.6.

O escopo da *Chemistry Education Research and Practice* é voltado para pesquisas e revisões de pesquisa em Ensino de Química; avaliações de práticas inovadoras eficazes no Ensino de Química; e análises aprofundadas de questões de relevância direta para o Ensino de Química. Os objetivos do periódico são: fornecer aos pesquisadores os meios para publicar seu trabalho na íntegra em um periódico dedicado exclusivamente ao Ensino de Química; oferecer aos professores de Química em todos os níveis um lugar onde ideias possam ser compartilhadas; fornecer métodos eficazes para o ensino e a aprendizagem de química; e aproximar professores e pesquisadores para que os resultados dos últimos sejam vistos pelos primeiros e que ambos os profissionais ganhem com essa troca de resultados.

Algo que vale ser destacado é que, coincidentemente, o editor da *Chemistry Education Research and Practice*, Scott Lewis, da *University of South Florida*, EUA, também aparece como um dos principais autores que publicaram artigos abordando a temática da Avaliação na área de Ensino de Química, como pode ser visto no Quadro 10.

Duas hipóteses são postas para reflexão, no que diz respeito ao porquê da revista *Chemistry Education Research and Practice* ter sido o principal periódico que publicou artigos sobre Avaliação na área de Ensino de Química. A

primeira refere-se a própria importância da revista e do seu impacto na área de Ensino de Química no contexto Internacional, atraindo o interesse de diversos pesquisadores para figurarem seus artigos na revista, aliado ao seu patrocínio do *Royal Society of Chemistry's Education Division*, que permite um alto volume de artigos publicados por ano em uma frequência igualmente alta anual de quatro volumes. A *Royal Society of Chemistry Education* é uma sociedade erudita, acadêmica e profissional do Reino Unido com o objetivo de “avançar as ciências químicas”, sendo formada em 1980 a partir da fusão da *Chemical Society*, do *Royal Institute of Chemistry*, da *Faraday Society* e da *Society for Analytical Chemistry*. Assim, o argumento poderia ser feito de que a revista em questão, a *Chemistry Education Research and Practice*, figuraria entre uma das revistas que mais publica artigos em Ensino de Química, independente do tema.

A segunda hipótese é a presença de um dos principais autores em Avaliação na área de Ensino de Química dos últimos 20 anos, presente como editor da revista. Seu interesse no tema como uma linha de pesquisa na área de Ensino de Química possa ter contribuído.

Segundo o Quadro 11, a outra revista que obteve um número considerável de artigos foi o *Journal of Chemical Education*. Ela é co-publicada pela *ACS Division of Chemical Education* e pela *ACS Publications*. Lançado em 1924, o *Journal of Chemical Education* é um dos principais periódicos de Ensino de Química do mundo. O periódico normalmente aborda conteúdo químico, experimentos de laboratório, métodos de ensino e pedagogias e serve como um meio de comunicação entre pessoas em todo o mundo, interessadas no ensino e na aprendizagem de Química. O fator de impacto de da revista nos últimos 5 anos é 2,7 e nos últimos 2 anos é 2,5. Similarmente a revista *Chemistry Education Research and Practice*, a revista *Journal of Chemical Education* possui seu editor, o Thomas A. Holme, como um dos autores que mais publicou artigos sobre Avaliação na área de Ensino de Química, como visto no Quadro 10. Mesmo que a revista tenha publicado vários artigos sobre Avaliação na área de Ensino de Químicao tema investigado neste estudo, sua representatividade no *corpus* não se compara a *Chemistry Education Research and Practice*, já que as revistas tiveram 16% e 66% dos artigos das últimas duas décadas, respectivamente.

Em relação aos países das publicações, apresentamos o Quadro 12

Quadro 12 - Países das publicações

Países	Nº de artigos
EUA	100
*Colaborações entre países	19
Turquia	16
Austrália	11
China	7
Reino Unido	7
Canadá	5
Alemanha	4
Indonésia	4
África do Sul	3
Grécia	3
Irlanda	3
Israel	3
Países Baixos	3
Espanha	2
Gana	2
República Sérvia	2
Singapura	2
Botswana	1
Finlândia	1
Índia	1
Malásia	1
Noruega	1
Paquistão	1
Polônia	1
Portugal	1
Suíça	1
Total	205

Fonte: os autores

Ao observar informações do Quadro 12, nota-se uma predominância de artigos norte-americanos no tema Avaliação na área de Ensino de Química, representando 49% do *corpus*. Entretanto, não se pode dizer que existe um interesse pelo tema concentrado apenas nesse país, já que foram identificadas pesquisas sobre Avaliação na área de Ensino de Química em 26 países no total. Outros países que merecem destacar em relação a sua representatividade são a Turquia com 8% das publicações totais e a Austrália com 5% das publicações. A quantidade de artigos que foram frutos de colaborações entre países também pode ser pontuada, compondo 9% do *corpus*.

Na sequência apresentamos o Quadro 13, em que identificamos as colaborações entre os diferentes países,

Quadro 13 - Colaborações entre Países

Países que colaboraram	Nº de artigos
Austrália & Croácia	1
China & Austrália	1
EUA & Alemanha	1
EUA & Suécia	1
EUA & Turquia	2
Israel & EUA	4
Malásia & Austrália	1
Países Baixos & EUA	1
Portugal & Austrália	1
República Checa, Turquia & Finlândia -	1
Singapura & Reino Unido	1
Suécia e Alemanha	1
Suécia, Austrália & EUA	1
Taiwan & Austrália	1
Turquia & Alemanha	1
Total	19

Fonte: os autores

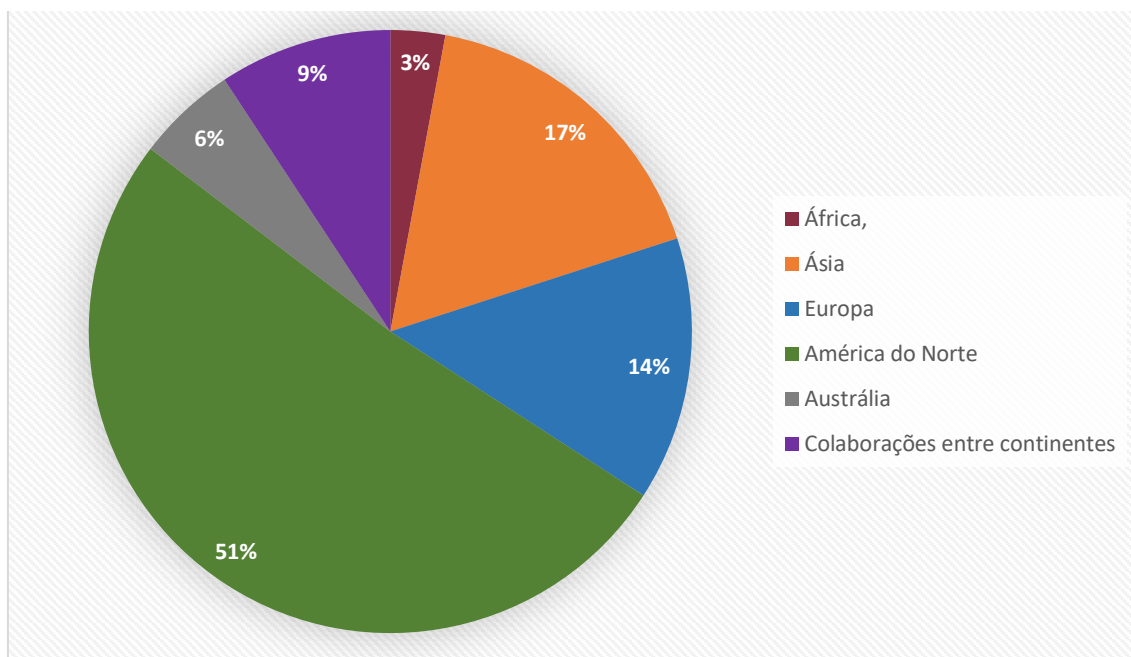
Para verificarmos a divisão dos artigos por continente e visualizarmos a concentração geográfica dos estudos, adotamos o modelo dos sete continentes, dividido em África, Ásia, Europa, América do Norte, América do Sul, Antártica e Austrália. Abaixo apresentamos esta distribuição:

Quadro 14 - Continentes das publicações

Continente	Nº de artigos
África	6
Ásia	35
Europa	29
América do Norte	105
Austrália	11
Colaborações entre continentes	19
Total	205

Fonte: os autores

Apresentamos abaixo a distribuição, em porcentagens, dos artigos por continente.



Fonte: os autores

A América do Norte se destaca em número de publicações sobre Avaliação na área de Ensino de Química com 51%, seguido da Ásia com 17%; Europa com 14%; Colaborações entre continentes com 9%; Austrália com 6% e África com 3%. Ao observarmos esses resultados, notamos a importância da Ásia nas publicações do contexto internacional, excedendo até mesmo a Europa em número de artigos sobre Avaliação na área de Ensino de Química.

Em relação aos níveis de ensino, concentramos nossas análises nos níveis de ensino que foram avaliados nos artigos. Como havia artigos de diferentes países, e conseqüentemente, diferentes sistemas educacionais, estes tiveram que ser padronizados para as nomenclaturas dos níveis educacionais da Educação Brasileira. Para isso, consultamos as descrições dos níveis educacionais e as idades dos alunos do *International Standard Classification of Education (ISCED)*²⁷ e tomamos como base a idade dos participantes das pesquisas relatadas pelos autores delas, quando o nível de ensino não era explicitado.

Quadro 15 – Níveis de Ensino

²⁷ A Classificação Internacional Padrão de Educação (ISCED) é uma estrutura desenvolvida pela UNESCO para categorizar e comparar programas educacionais e qualificações entre países. Estabelecida para fornecer uma maneira sistemática de coletar e analisar estatísticas educacionais, a ISCED ajuda a entender a estrutura dos sistemas educacionais em todo o mundo.

Nº de artigos	Categoria	Subcategoria
8	Professores e alunos	Professores de Universidade e alunos da graduação [4]
		Professores de Universidade e alunos da pós-graduação [2]
		Professores e licenciandos [2]
102	Ensino Superior ²⁸	Faculdade [8]
		Graduação [41]
		Faculdade e Universidade [1]
		Universidade [44]
		Ensino Superior [3]
		Pós Graduação [4]
		Pré-Graduação [1]
36	Educação Básica	Ensino Fundamental Anos Finais [6]
		Ensino Fundamental e Ensino Médio [1]
		Ensino Fundamental Anos Finais e Ensino Médio [10]
		Ensino Médio [19]
6	Educação Básica e Ensino Superior	Ensino Médio e Pré-Graduação [1]
		Ensino Médio e Graduação [1]
		Ensino Médio e Universidade [1]
		Ensino Médio e Licenciandos [1]
		Ensino Médio e Faculdade [1]
		Educação Básica e Faculdade [1]
12	Licenciandos	-
19	Professores	Professores em serviço [12]
		Corpo Docente do Ensino Superior [2]
		Corpo Docente de Universidade [4]
		Corpo Docente de Universidade e professores [1]
3	Multiníveis ²⁹	Ensino Médio, Graduação, Pós-Graduação, Docentes da faculdade [1]
		Ensino Fundamental Anos Finais, Ensino Médio e professores [1]
		Graduação, Licenciandos e Professores em serviço [1]
18	Não especificado	-

Fonte: os autores

A terceira coluna do Quadro15 fornece informações mais detalhadas a respeito do nível de ensino avaliado nos artigos, podendo ser útil para pesquisadores que desejam buscar pesquisas sobre Avaliação na área de Ensino de Química com

²⁸ Considerou-se como Ensino Superior, “faculdade”, “Educação Terciária”, “Graduação” e “Universidade”.

²⁹ Artigos que investigaram 3 níveis de ensino ou mais foram inseridos nesta categoria.

sujeitos específicos. Entretanto a primeira e segunda coluna revelam tendências de pesquisa mais abrangentes sobre a literatura sobre Avaliação na área de Ensino de Química, das quais serão discutidos na sequência.

No que diz respeito ao nível, nota-se que a principal tendência entre os artigos sobre Avaliação na área de Ensino de Química é avaliar sujeitos do Ensino Superior, já que 50% dos artigos apresentaram essa característica (102 artigos). Na sequência, também observamos a tendência, porém em menor grau, de avaliar sujeitos da Educação Básica, considerando que 8% dos artigos apresentaram essa característica (36 artigos).

Outra característica do *corpus* é sobre a presença de artigos que avaliam professores. 8 artigos avaliaram professores e alunos; 19 artigos avaliaram apenas professores; e 3 avaliaram multiníveis (sendo que todos os artigos desse agrupamento incluíram professores, além de outros níveis). Logo, 30 artigos, ou seja, 15% do *corpus*. Esse é um achado interessante, pois, convencionalmente, quando se pensa em avaliação, a primeira associação que fazemos é relacioná-la a avaliação de alunos, porém os resultados destacam a presença de artigos que avaliam professores também. Dado que o *corpus* é elevado, é normal haver uma maior divisão dos agrupamentos, porém a porcentagem de artigos que avaliam professores é maior do que inicialmente esperávamos antes de realizarmos a revisão.

Em relação às instituições das publicações, identificamos as com maior quantidade de artigos, apresentadas no Quadro 16.

Quadro 16 - Instituições com maior quantidade de artigos

Nº de artigos	Instituição
16	University of South Florida, Estados Unidos
10	Miami University, Estados Unidos
9	University of Wisconsin, Estados Unidos
6	Middle East Technical University, Turquia
6	Purdue University, Estados Unidos
5	Technion-Israel Institute of Technology, Israel
5	University of Ottawa, Canada
5	University of California, Estados Unidos

Fonte: os autores

Também foi realizado um levantamento a respeito das outras instituições presentes no *corpus*, porém com menor número de artigos. Identificamos as instituições apresentadas no Quadro 17.

Quadro 17 - Instituições com menor quantidade de artigos

Nº de artigos	Instituição
4	University of Nebraska-Lincoln, Estados Unidos Clemson University, Estados Unidos Van Yuzuncu Yil University, Turquia Michigan State University, Estados Unidos Berkeley University of California, Estados Unidos California State University, Estados Unidos
3	University of Kiel – Alemanha University of Haifa – Israel Curtin University – Australia Portland State University – Estados Unidos University of Minnesota – Estados Unidos University of Michigan – Estados Unidos University of Central Florida – Estados Unidos Iowa State University – Estados Unidos University of Iowa – Estados Unidos North Carolina State University – Estados Unidos
2	University of Novi Sad – Sérvia University of New South Wales – Australia University of Macau – China James Madison University – Estados Unidos Dokuz Eylul University – Turquia The IMMEX Project – Estados Unidos Deakin University – Australia University of Bremen – Alemanha University of Sydney – Australia North Dakota State University- Estados Unidos Leiden University – Países Baixos National and Kapodistrian University of Athens – Grécia Samuel Neaman Institute – Israel Massachusetts Institute of Technology – Estados Unidos University of Massachusetts – Estados Unidos Utrecht university – Países Baixos Indiana University – Estados Unidos Seattle University – Estados Unidos Uppsala University – Suécia University of Arizona – Estados Unidos Queensland university of technology – Australia Nanyang technological university – Singapura University of LimERICk – Irlanda Glendale Community College – Estados Unidos University of Edinburgh – Escócia Metropolitan State University – Estados Unidos Georgia institute of technology – Estados Unidos University of North Carolina – Estados Unidos University of Education, Winneba – Gana

Fonte: os autores

Ao analisar o Quadro 16, nota-se que *University of South Florida*, Estados Unidos, foi a instituição com maior quantidade de artigos, com 16 artigos, representando 8% do corpus. Dentre as instituições identificadas no Quadro 16, cinco são dos Estados Unidos: *University of South Florida*, *Miami University*, *University of Wisconsin*, *Purdue University* e a *University of California*. Outras instituições que figuram entre as mais produtivas em Avaliação na área de Ensino de Química foram: a *Middle East Technical University* da Turquia com 6 artigos; *Technion-Israel Institute of Technology* de Israel com 5 artigos; e *University of Ottawa* do Canada com 5 artigos.

Similarmente a prevalência de pesquisas originadas na Flórida, ao identificarmos os principais autores, esse fenômeno também ocorre ao olharmos para as instituições. A *University of South Florida* e a *Miami University* são as duas instituições com mais artigos do *corpus* e estão localizadas na Flórida, nos Estados Unidos.

4.1.1 UMA SÍNTESE CRÍTICA DAS TENDÊNCIAS E LACUNAS - PARTE I

O aumento no número de publicações sobre a temática avaliação na área de ensino de Química, especialmente entre 2011 e 2021, reflete uma evolução significativa nas práticas educacionais e nas prioridades da comunidade acadêmica. Durante a primeira década analisada (2002-2011), apenas 18% dos artigos Avaliação na área de Ensino de Química foram publicados, enquanto na segunda década (2012-2021) esse número aumentou significativamente para 82%. Essa mudança acentuada reflete uma possível valorização da avaliação como um campo de pesquisa, bem como a possibilidade de discuti-la em diferentes níveis de ensino e espaços formativos.

Acreditamos que a introdução e a popularização do conceito de avaliação formativa, especialmente após os trabalhos de Black e Wiliam em 1998, também desempenharam um papel relevante no aumento do interesse pelo tema. As avaliações formativas são projetadas para monitorar o progresso dos alunos durante o processo de aprendizagem, permitindo ajustes imediatos nas práticas de ensino. A abordagem não apenas ajuda os educadores a responderem às necessidades dos alunos de forma mais eficaz, mas também promove um ambiente de aprendizagem mais dinâmico e responsivo, centrado no aluno, se adequando melhor ao ensino centrado no aluno. Acreditamos que essa mudança de paradigma no ensino, que

ênfatisa a participação ativa dos estudantes em seu próprio aprendizado, tem se espalhado por todo o setor educacional, refletindo-se nas publicações sobre avaliação.

Além disso, a maioria dos artigos foi publicada em dois periódicos: *Chemistry Education Research and Practice* e o *Journal of Chemical Education*, com 66% e 16% das publicações, respectivamente. Assim, esses periódicos se caracterizam como plataformas de disseminação do conhecimento fundamentais para a divulgação de pesquisas em avaliação na área de Ensino de Química e oferecem um espaço para discussões relevantes no que diz respeito às práticas avaliativas adotadas pela comunidade de professores e pesquisadores em Ensino de Química.

A concentração de autores e instituições nos Estados Unidos, especialmente na Flórida, é outro aspecto notável das publicações investigadas. Avaliação na área de Ensino de Química A *University of South Florida* e a *Miami University* destacam-se por sua alta produtividade. Avaliação na área de Ensino de Química Esse fenômeno pode ser atribuído à presença de pesquisadores influentes nessas instituições, que têm contribuído significativamente para a literatura da área, já que os três autores mais produtivos (Ellen J. Yeziarski, Scott E. Lewis e Jennifer E. Lewis) pertencem a universidades no estado da Flórida.

Em termos de nível de ensino, a maioria dos artigos focou no ensino superior, representando 50% do total, enquanto a educação básica correspondeu a apenas 8%. Essa predominância do ensino superior pode estar relacionada à complexidade e ao rigor das avaliações necessárias nesse nível educacional, onde as expectativas de aprendizagem são mais elevadas. Além disso, uma parte considerável dos artigos também investigou a avaliação de professores, refletindo uma perspectiva mais ampla sobre a avaliação no contexto educacional, que considera outros agentes envolvidos no contexto escolar. Consideramos esse resultado interessante, pois comumente a avaliação da aprendizagem, focada nos estudantes, é priorizada nas discussões a respeito de avaliação, porém a atenção para aspectos da avaliação e à docência indicam interfaces importantes para os processos educacionais.

4.2 O QUE TEM SIDO AVALIADO NOS ARTIGOS INTERNACIONAIS SOBRE AVALIAÇÃO NA ÁREA DE ENSINO DE QUÍMICA?

Em relação ao que foi avaliado nos artigos sobre Avaliação na área de Ensino de Química, foram identificadas 10 categorias que acomodaram todos os artigos analisados e seguem apresentadas no Quadro 18. Para realizar a categorização, ao abrirmos os artigos, pressionamos “Ctrl + f” e escrevemos “*assess*” para destacar todos os excertos que continham os termos “*assess*” e suas diversas conjugações e derivações, como por exemplo: “*assess, assesses, assessed, assessing e assessment*”. Isso nos permitiu analisar todas as frases com a palavra e compreender o que de fato foi avaliado em cada artigo. É importante ressaltar que as categorizações foram validadas por uma segunda pesquisadora para garantir a confiabilidade dos resultados. Destacamos que as categorizações não foram realizadas em relação ao foco central dos artigos, mas sim a respeito do que foi de fato avaliado.

Quadro 18 - Resumo das categorizações

Categoria	Nº de artigos	Porcentagem
C01: Avalia habilidades	18	8,8%
C02: Avalia estratégias de avaliação ³⁰ , provas, exames e itens	17	8,3%
C03: Avalia propostas de ensino ³¹	39	19,0%
C04: Avalia um instrumento, ferramenta ou tecnologia ³²	36	17,5%
C05: Avalia compreensões, percepções e interesse	19	9,3%
C06: Avalia conhecimentos	20	9,8%
C07: Avalia o currículo	14	6,8%
C08: Avalia atitudes	10	4,9%
C09: Avalia outras pesquisas ³³	3	1,5%
C10: Discute avaliação em contextos diferentes	29	14,1%
Total	205	100%

Fonte: os autores

Ao analisarmos o Quadro 18, observa-se que a categoria mais representativa foi a C03: Avalia propostas de ensino, representando 19% do *corpus*. De fato, isso foi uma grande tendência no *corpus*. Os autores dos artigos analisados

³⁰ Traduzido do Inglês *assessments*. Vários artigos utilizam o termo *assessments*. *Assessment*, em inglês, possui um sentido de prova mais formal, as vezes aplicada em larga escala, contrastado com o termo *test*, que seria a forma mais casual de prova.

³¹ Consideramos como proposta de ensino; abordagem, tarefa, intervenção, framework e workshop.

³² Todos os instrumentos ou ferramentas que envolvem mídia, foram deslocados para a subcategoria de tecnologias.

³³ As categorias indicam o que foi avaliado. Na categoria C09 alocamos os únicos artigos que *avaliaram outros artigos*.

apresentaram um grande interesse em avaliar a qualidade, o impacto, a importância, o efeito, características, entre outros aspectos, de várias propostas de ensino. Essas propostas incluíram uma grande diversidade de abordagens, tarefas, intervenções, frameworks³⁴, atividades, workshops, entre outros.

Em seguida, observa-se uma tendência em avaliar instrumentos, ferramentas ou tecnologias, evidenciada pela alocação de 17,5% dos artigos na categoria C04: Avalia um instrumento, ferramenta ou tecnologia.

Identificamos também um conjunto de artigos que não avaliaram algo, propriamente dito, mas apresentaram discussões teóricas a respeito de aspectos da avaliação em um contexto específico. Essa categoria se distingue das outras, por esse motivo. Isso é evidente devido a 14,1% dos artigos terem sido alocados na categoria C10: Discute avaliação em contextos diferentes.

No Quadro 19 apresentamos as definições de cada uma das categorias. Essas conceituações foram pautadas nas análises realizadas dos artigos do *corpus*.

Quadro 19 - Definições das categorias

Categoria	Definição
C01: Avalia habilidades	A categoria C01 engloba artigos que avaliaram habilidades. Alguns exemplos foram habilidades de: pensamento crítico, abstração; raciocínio gráfico; laboratório; pensamento científico; resolução de problemas; leitura, etc. ³⁵
C02: Avalia estratégias de avaliação, provas, exames e itens	A categoria C02 acomoda artigos que avaliaram estratégias de avaliação, provas, exames e itens. Alguns exemplos de artigos incluídos nesta categoria foram: artigos que avaliam exames de um módulo introdutório de química, avaliações abertas, avaliações multimodais, estratégias de avaliação baseado em competências, etc. ³⁶
C03: Avalia propostas de ensino	A categoria C03 reúne artigos que avaliaram propostas de ensino. Alguns exemplos de propostas de ensino que foram avaliadas foram: abordagens, tarefas, intervenções, frameworks, workshops, atividades, sequências didáticas, etc. ³⁷
C04: Avalia um instrumento,	A categoria C04 reúne artigos que avaliaram um instrumento, ferramenta ou tecnologia. Alguns exemplos

³⁴ Na educação, um “framework” é uma estrutura conceitual em que princípios organizadores constroem uma visão comum. A visão então orienta o desenvolvimento do currículo e/ou instrução, incluindo experiências de campo (Collins; O'Brien, 2003).

³⁵ A lista completa de habilidades avaliadas pela categoria C01 pode ser vista na subsubseção 4.2.1.

³⁶ A lista completa de estratégias de avaliação, provas, exames e itens avaliadas pela categoria C02 pode ser vista na subsubseção 4.2.2.

³⁷ A lista completa de propostas de ensino avaliadas pela categoria C03 pode ser vista na subsubseção 4.2.3.

ferramenta ou tecnologia	foram: mapas conceituais, ferramentas de medição de habilidades, ambientes virtuais, simulações, softwares, etc. ³⁸
C05: Avalia compreensões, percepções e interesse	A categoria C05 reúne artigos que avaliaram compreensões, percepções e interesse. Alguns exemplos do que foi avaliado foram compreensões e percepções a respeito de: fenômenos químicos, reações químicas, representações químicas, estruturas químicas, etc. ³⁹
C06: Avalia conhecimentos	A categoria C06 reúne artigos que avaliaram conhecimentos. Alguns exemplos dos conhecimentos avaliados foram a respeito de: conceitos básicos da química, conhecimentos algoritmos, conhecimentos computacionais, etc. ⁴⁰
C07: Avalia o currículo	A categoria C07 reúne artigos que avaliaram o currículo. Alguns exemplos dos currículos avaliados foram: documentos estruturadores para o curso de bacharelado em química, objetivos educacionais presente no currículo, currículos de disciplinas da graduação, programas de formação de professores, etc. ⁴¹
C08: Avalia atitudes	A categoria C08 reúne artigos que avaliaram atitudes. Alguns exemplos das atitudes avaliadas foram: atitudes emocionais, atitudes intelectuais, atitudes positivas, práticas de avaliação, etc. ⁴²
C09: Avalia outras pesquisas	A categoria C09 reúne artigos que avaliaram pesquisas. Alguns exemplos das pesquisas foram: artigos sobre avaliação na pandemia, pesquisas que usaram avaliações quantitativas, a relação entre pesquisa e ensino, etc. ⁴³
C10: Discute avaliação em contextos diferentes	A categoria C10 reúne artigos que discutiram a avaliação em um contexto específico. Alguns exemplos dos contextos discutidos foram: a avaliação em propostas de ensino, a avaliação na perspectiva do aluno ou professor, tecnologia e avaliação, etc. ⁴⁴

Fonte: os autores

Retornando ao movimento analítico empregado para realizar as categorizações, utilizamos os seguintes dicionários: *Houaiss* e o *The Greenwood Dictionary of Education* para definirmos os objetos avaliados, contidos nos títulos das categorias do Quadro 18. Apresentamos as definições utilizadas no Quadro 20.

³⁸ A lista completa de instrumentos, ferramentas ou tecnologias avaliadas pela categoria C04 pode ser vista na subsubseção 4.2.4.

³⁹ A lista completa de compreensões, percepções e interesse avaliados pela categoria C05 pode ser vista na subsubseção 4.2.5.

⁴⁰ A lista completa conhecimentos avaliados pela categoria C06 pode ser vista na subsubseção 4.2.6.

⁴¹ A lista completa de currículos avaliados pela categoria C07 pode ser vista na subsubseção 4.2.7.

⁴² A lista completa de atitudes avaliadas pela categoria C08 pode ser vista na subsubseção 4.2.8.

⁴³ A lista completa de pesquisas avaliadas pela categoria C09 pode ser vista na subsubseção 4.2.9.

⁴⁴ A lista completa de discussões realizadas pela categoria C10 pode ser vista na subsubseção 4.2.10.

Quadro 20 - Definições utilizadas para realizar as categorizações

Habilidade
1. Qualidade ou característica de quem é hábil; 2. Algo que se sabe fazer. Na educação, a habilidade geralmente está relacionada à leitura, adição, direção, contabilidade, etc. A habilidade é alcançada pela prática e pode ter muitos componentes. 3. Capacidade de executar uma tarefa, como em uma "alta habilidade" ou "baixa habilidade". Grau de habilidade no desempenho da tarefa.
Avaliação
1. Cálculo do valor de um bem ou bens 2. Valor determinado por quem avalia 3. Apreciação ou conjectura sobre condições, extensão, intensidade, qualidade de algo 4. Verificação que objetiva determinar a competência, o progresso, etc. de um profissional, aluno etc. 5. Qualquer método usado para entender melhor o conhecimento atual que um aluno possui. A avaliação pode afetar as decisões sobre notas, avanço, colocação, necessidades de instrução e currículo. Informações coletadas podem incluir atividades sociais, educacionais, e observações psicológicas usadas para identificar os pontos fortes e fracos de um indivíduo. Os métodos e procedimentos usados em coleta e interpretação de informações sobre estudantes, instituições e programas para fins de avaliação, apreciação, etc.
Prova/exame
1. Trabalho escolar geralmente composto de uma série de perguntas, que tem por finalidade avaliar os conhecimentos do aluno; teste; exame. 2. Um exercício usado para avaliar conhecimento ou habilidade. Frequentemente, um conjunto formal de perguntas ou tarefas destinadas a gerar uma representação quantitativa usada para determinar se um aluno possui certas habilidades ou compreende informações dadas. Os testes podem ser usados para comparar indivíduos a grupos ou populações ou pode ser usado para avaliar o desenvolvimento individual.
Propostas de Ensino
1. Consideramos como proposta de ensino; abordagem, tarefa, intervenção, framework e workshop.
Instrumento
1. Objeto simples ou constituído por várias peças, que serve para executar um trabalho, fazer uma medição ou observação etc.
Ferramenta
1. Qualquer instrumento necessário à prática profissional. 2. Meio para alcançar um fim.
Tecnologia
1. Qualquer técnica moderna e complexa
Compreensão
1. Faculdade de entender, de perceber o significado de algo; entendimento. 2. Perfeito domínio intelectual de um assunto
Conhecimento
1. Ato ou efeito de conhecer. 2. Faculdade de conhecer 3. Fato ou condição de estar ciente ou consciente de algo; ciência, informação, notícia.

4. Somatório do que se conhece; conjunto das informações e princípios armazenados pela humanidade
Percepção
1. Faculdade de apreender por meio dos sentidos ou da mente. 2. Consciência (de alguma coisa ou pessoa), impressão ou intuição, especialmente moral.
Currículo
1. Programação total ou parcial de um curso ou de uma matéria a ser examinada. 2. As definições de currículo são tão numerosas quanto há filósofos da educação. O modelo de currículo de Tyler traz as necessidades da criança, a sociedade e as disciplinas acadêmicas juntas na tentativa de formular os objetivos gerais que os professores devem ensinar e os alunos devem aprender. 3. Para J. Allen Queen, o currículo é como uma exigência da sociedade do processo educacional formal preparar as próximas gerações para adaptabilidade, aceitação, diversidade e sobrevivência em um mundo desconhecido. 4. O currículo também pode ser visto como uma coleção de processos formais e processos informais e experiências que preparam os alunos para alcançar e se adaptar às necessidades pessoais e sociais. 5. Os cursos de estudo oferecido em uma instituição de ensino representam o currículo aberto e intencional, o que tornamos público e anunciamos que somos ensinados na escola. O currículo aberto ou explícito é o que colocamos em nossos catálogos de cursos e em nossos programas de cursos. 6. Há também um currículo oculto ao que ensinamos que não divulgamos, e muitas vezes nem mesmo estamos conscientes de que estamos ensinando, ainda assim é ensinado. 7. Currículo em seu sentido mais amplo significa o que ensinamos.
Atitudes
1. Comportamento ditado por disposição interior, maneira, conduta 2. Existem na mente de uma pessoa como abstrações e não pode ser quantificado. Essas crenças representam uma disposição em relação a pessoas, comportamentos ou coisas. Podem ser observadas ações que são feitas com base nessas crenças. Lei limita algumas destas, algumas são vinculados pela sociedade (família e amigos). Os valores moldam as atitudes de um indivíduo em relação a ações, pessoas e coisas. As atitudes também direcionam as aspirações de uma pessoa e ambições,

Fonte: os autores

Assim, as definições apresentadas no Quadro 20, em conjunto com os termos que os próprios autores utilizaram em seus artigos, nos auxiliaram nos momentos de categorização a fim de assegurar clareza nesse movimento.

Na sequência, apresentamos em detalhes as 10 categorias estabelecidas para a análise dos artigos. Apresentamos também as subcategorias, em que visamos explicitar o que foi avaliado em cada um dos 205 artigos. Também

incluiremos citações diretas de alguns artigos selecionados de cada categoria para exemplificar cada uma das categorias, auxiliando-nos nas discussões.

4.2.1 Avalia habilidades (C01)

A categoria C01 consistiu dos artigos que avaliaram habilidades. No Quadro 21 apresentamos todas as subcategorias encontradas na categoria C01 e os respectivos códigos desses artigos. As subcategorias mostram quais habilidades específicas foram avaliadas, exatamente.

Quadro 21 - C01: Avalia habilidades

Subcategoria	Código
Avalia o pensamento crítico	A001
Avalia habilidades de transferência de conhecimento	A005
Avalia habilidades de processo científico	A015
Avalia a habilidade de abstração dos alunos na Química Orgânica	A033
Avalia o raciocínio gráfico	A047
Avalia a habilidade de linkar conceitos prévios a novos conceitos	A048
Avalia habilidades de laboratório	A056
Avalia habilidades de pensamento científico	A080
Avalia habilidades de pensamento de sistemas de ordem mais alta	A099
Avalia a autoavaliação como expressão de habilidade metacognitiva e a aprendizagem em estequiometria	A121
Avalia o “student thinking”, ou seja, o raciocínio do estudante. Buscou entender como os alunos trabalharam em duas reações ácido-base usando papel-lápis ou um aplicativo.	A070
Avalia estratégias de aprendizagem dos alunos, aprendizagem conceitual e resolução de problemas	A130
Avalia a capacidade de leitura e as habilidades cognitivas na dificuldade do item e a avaliação do modelo de competência	A155
Avalia a habilidade dos alunos de balancear equações químicas e desenhar representações particuladas apropriadas dessas reações, por meio de uma ferramenta de desenho	A160
Avalia a habilidade do professor em integrar os componentes do conhecimento pedagógico do conteúdo em aulas	A170
Avalia a habilidade dos alunos de coordenar três níveis de representação em química	A178
Avalia habilidades algorítmicas, habilidades cognitivas de ordem inferior e habilidades cognitivas de ordem superior	A197
Avalia as habilidades cognitivas de ordem superior, habilidades eficazes de comunicação científica e a compreensão da natureza da ciência	A120

Fonte: os autores

Ao analisar o Quadro 21, nota-se uma grande variedade de habilidades que foram avaliadas nos artigos. No que diz respeito aos sujeitos, apenas

o A170 da categoria C1 avaliou a habilidade do professor. Os demais artigos avaliaram habilidades dos alunos.

No geral, as habilidades avaliadas na categoria C01 foram: pensamento crítico; transferência de conhecimento; processo científico; abstração; raciocínio gráfico; conexão de conceitos; habilidades de laboratório; pensamento científico; pensamento de sistemas de ordem mais alta; autoavaliação; raciocínio; resolução de problemas; leitura; habilidades cognitivas; balanceamento de equações; integração de componentes do conhecimento pedagógico do conteúdo; coordenação de níveis de representação; habilidades algorítmicas; habilidades cognitivas de ordem inferior; habilidades cognitivas de ordem superior; e comunicação científica.

As alocações da categoria C01 puderam ser realizadas muitas vezes com base na própria palavra “habilidades” frequentemente utilizada pelos autores dos artigos. Quando esse termo não foi utilizado no artigo, recorremos a definição do que era avaliado, que continha o termo habilidade explicitamente ou implicitamente.

Trazemos abaixo um excerto do A015 para exemplificar a categoria C01:

A pesquisa foi implementada relacionada à análise das **habilidades de processo científico dos alunos** por meio do aprendizado baseado em projetos na era da pandemia da Covid-19. [...] Os instrumentos usados neste estudo foram a folha de avaliação de habilidades de processo científico, a folha de observação de atividades e a resposta dos alunos ao questionário. Houve duas habilidades mais altas mostradas pelos alunos, a saber, (1) determinar ferramentas e materiais e (2) determinar indicadores de variáveis de pesquisa. As duas habilidades estavam na categoria excelente. Habilidades para determinar as etapas do trabalho e fazer uma tabela de dados estavam na categoria boa. A categoria muito boa conteve as habilidades de formular objetivos de pesquisa, fazer hipóteses, analisar e tirar conclusões. Fazer uma declaração de problema estava na categoria ruim. **Em geral, as habilidades de processo científico dos alunos estavam na categoria boa.** A atividade de implementação de projetos estudantis recebeu uma categoria excelente e os alunos deram uma resposta positiva à implementação do projeto durante uma pandemia (A015, p. 371, grifo nosso, tradução nossa)⁴⁵.

⁴⁵ No texto original: Research has been implemented related to the analysis of the students' science process skills through project-based learning in a Covid-19 pandemic era. The prohibition of face-to-face lectures causes hindered experimental activities to be conducted on campus. This condition encourages project-based learning to carry out extraction experiments independently (Self-Project Based Learning, SjBL). The method used in this study was the Pre-Experimental Design One-Shot Case Study involving 94 4th semester chemistry education students. The instruments used in this study were the science process skills assessment sheet, the activity observation sheet, and the students' response of the questionnaire. There were two highest skills showed by the students, namely, (1) determining

Observamos que o A015 avaliou as habilidades de processo científico de alunos do quarto semestre de um curso de licenciatura em química. Essas habilidades são definidas pelo A015 como: “a capacidade de aplicar métodos científicos para construir conhecimento, entender, desenvolver e descobrir ciência para resolver problemas e formular resultados” (Rusmini; Suyono; Agustini, 2021, p. 372, tradução nossa)⁴⁶. As habilidades de processo científico podem ser divididas em básicas e integradas. Alguns exemplos das habilidades básicas são: observar, usar relações de tempo ou espaço, inferir, medir, comunicar, classificar e prever. Alguns exemplos de habilidade consideradas integradas são: controlar variáveis, definir operacionalmente, formular hipóteses, interpretar dados, experimentar, formular modelos e apresentar informações.

De acordo com o A015, o método utilizado no artigo foi o “Estudo de Caso Único de Design Pré-Experimental”, em que categorizaram o domínio das habilidades do processo científico em categorias determinadas como “excelente”, “muito bom”, “bom” ou “ruim”, por exemplo. A aprendizagem baseada em projetos implementada no A015 foi favorável ao desenvolvimento das habilidades do processo científico, entretanto, para o A015 ainda são necessários esforços para desenvolver habilidades de processo científico para futuros professores de química, uma vez que essa é uma preocupação recente na formação inicial.

Trazemos abaixo um excerto do A033 para melhor ilustrar a categoria C01:

Os alunos frequentemente têm dificuldades para resolver problemas de mecanismo em cursos de química orgânica. Eles frequentemente focam em características de superfície, têm dificuldade em atribuir significado a símbolos e não reconhecem tarefas que são diferentes das tarefas exatas praticadas. Para ter mais sucesso, os alunos

tools and materials and (2) determining research variables indicators. The two skills were in excellent category. Skills to determine the work steps, and to make a data table including good categories. Quite good category was in the skills of making research objectives, making hypotheses, analyzing and drawing conclusions. Making a problem statement was in the bad category. In general, students' science process skills were in the good category. The activity of implementing student projects received an excellent category and students gave a positive response to project implementation during a pandemic. The results of this study contribute to science learning in the future. Efforts are needed to train science process skills to prospective chemistry teachers so that teachers who have good science process skills are produced (A015, 2021, p. 371).

⁴⁶ No texto original: Science process skills (SPS) were the ability to apply scientific methods to construct knowledge, understand, develop and discover science in order to solve problems and formulate results [...] (A015, 2021, p. 372).

precisam ser capazes de extrair características, mapear similaridades com problemas vistos anteriormente e extrapolar enquanto resolvem problemas. Em suma, os alunos devem ser capazes de reconhecer e gerar abstrações. [...] aplicamos o modelo de mapeamento de representação de Hahn e Chater (1998^a) para caracterizar a abstração empregada pelos alunos ao resolver problemas mecanicistas em química orgânica e para medir o crescimento dos alunos na capacidade de abstração ao longo de um semestre. Este modelo operacionaliza a abstração considerando (a) as maneiras pelas quais os alunos combinam o conhecimento existente com novas instâncias (abstração) e (b) o nível de abstração das representações dos alunos (A033, 2017, p. 169, tradução nossa)⁴⁷.

Nota-se que o A033 avalia a habilidade de abstração de alunos em problemas de química orgânica. De acordo com o A033, em química orgânica, é necessário a utilização de conceitos abstratos e visualizações para resolver problemas. Apesar disso, os alunos frequentemente possuem dificuldades nesses aspectos, como por exemplo, ao desenharem mecanismos de reações. Isso ocorre, pois os alunos não possuem uma compreensão profunda dos símbolos que usam, levando a uma abordagem superficial. Essas manifestações dos alunos, são decorrentes de uma memorização, que não envolve raciocínio, dificultando que os alunos compreendam os princípios por trás dos mecanismos.

Para melhorar as habilidades de resolução de problemas dos alunos, os professores precisam compreender como os alunos desenvolvem sua capacidade de abstração. Para isso, o A033 utilizou a estrutura de mapeamento de representação de Hahn e Chater (1998) para avaliar as habilidades de abstração dos alunos e informar estratégias de ensino.

A estrutura supracitada auxilia na categorização de tipos de raciocínio e enfatiza a interação entre representações abstratas e concretas. Ao contrário de outras estruturas que assumem um corpo fixo de conhecimento, o modelo de

⁴⁷ No texto original: "Students often struggle with solving mechanism problems in organic chemistry courses. They frequently focus on surface features, have difficulty attributing meaning to symbols, and do not recognize tasks that are different from the exact tasks practiced. To be more successful, students need to be able to extract salient features, map similarities to problems seen previously, and extrapolate while solving problems. In short, students must be able to recognize and generate abstractions. To help students in learning to solve problems, we need a better understanding of the nature of students' capacity for abstraction. Building upon an exploratory study (Sevian H., Bernholt S., Szeinberg G. A., Auguste S. and Pérez L. C., (2015), Use of representation mapping to capture abstraction in problem solving in different courses in chemistry, Chem. Educ. Res. Pract., 16(3), 429–446), we applied the representation mapping model of Hahn and Chater (1998a) to characterize the abstraction employed by students while solving mechanistic problems in organic chemistry, and to measure students' growth in abstraction capacity across a semester. This model operationalizes abstraction by considering (a) the ways in which students match existing knowledge to new instances (abstracting) and (b) the level of abstractness of students' representations" (Weinrich; Sevian, 2017, p. 169).

mapeamento de representação de Hahn e Chater (1998) captura a natureza dinâmica dos processos cognitivos dos alunos. A estrutura define dois tipos principais de representações: representação de nova instância, que reflete como um aluno percebe e aborda um problema específico, e conhecimento armazenado, que consiste nos recursos aplicados para resolvê-lo.

A respeito do funcionamento da estrutura, para o A033, este avalia a abstração de duas formas: abstração (substantivo) e abstração (verbo). A abstração (substantivo): refere-se ao nível de abstração na representação mental de um aluno, indicando o quão distante ele está de características concretas. Por exemplo, usar símbolos numéricos para descrever um problema é considerado mais abstrato do que usar gestos com as mãos. Já a abstração (verbo): descreve a ação de mover-se entre representações, mostrando como os alunos conectam seu conhecimento armazenado ao novo contexto do problema.

Ao comparar a abstração das representações, a estruturação pode identificar a eficácia com que os alunos utilizam o conhecimento prévio. Por exemplo, um aluno focado em características explícitas de um problema demonstra menor abstração, enquanto outro que conecta conceitos ou princípios subjacentes demonstra maior abstração.

A respeito da metodologia do A033, o estudo foi conduzido em uma universidade no nordeste dos EUA em um curso de Química Orgânica II. 20 alunos foram recrutados para entrevistas. As entrevistas foram semiestruturadas e exploraram os processos de resolução de problemas dos alunos em questões de prova. O professor identificou o material-chave e selecionou problemas representativos antes de dois exames. Os alunos foram entrevistados após os exames, com foco em suas abordagens para problemas de mecanismo na exposição de suas etapas de resolução de problemas ao responderem às perguntas.

A respeito dos resultados do A033, quatro dimensões da abstração nas representações dos alunos foram identificadas. Estas foram: características explícitas vs. implícitas, sequência de eventos vs. explicações, estrutura vs. função e terminologia específica vs. generalizada. A maioria dos alunos se concentrou mais na estrutura do que na função, e os alunos de alto desempenho foram melhores em explicar suas abstrações. O A033 também relata um aumento na abstração dos alunos ao longo do semestre. Isso foi evidenciado por mudanças perceptíveis nos tipos de raciocínio utilizados pelos alunos, saindo de um raciocínio baseado na

memória para um raciocínio mais abstrato, o que demonstrou uma compreensão mais clara dos mecanismos de reação.

A respeito de recomendações para estudos posteriores, o A033 sugere que as avaliações devem ser elaboradas para encorajar múltiplos processos de raciocínio e tipos de abstração. Assim, os professores devem criar problemas que permitam a prática com várias estratégias de raciocínio e devem incluir momentos de feedback aos alunos, que mostrem múltiplas abordagens de solução. Essa interação pode provocar diferentes tipos de abstração nos alunos e é necessária para desenvolver a flexibilidade dos alunos na resolução de problemas.

4.2.2 Avalia estratégias de avaliação, provas, exames e itens (C02)

A categoria C02 consistiu dos artigos que avaliaram estratégias de avaliação, provas, exames e itens. No Quadro 22 apresentamos todas as subcategorias encontradas na categoria C02 e os respectivos códigos desses artigos. As subcategorias mostram quais avaliações, provas, exames e itens específicos foram avaliadas, exatamente.

Quadro 22 - C02: Avalia estratégias de avaliação, provas, exames e itens

Subcategoria	Código
Avalia exames de química, para jovens de 15 a 16 anos na Inglaterra de uma nova pedagogia de trabalho prático	A002
Avalia um exame de um módulo introdutório de química, usando a dimensão semântica.	A021
Avalia tipos de avaliação usados no ensino de química pós-secundário	A026
Avalia uma avaliação aberta específica, denominada Exercícios Criativos (EC)	A077
Avalia um exame de química geral do primeiro período do ACS <i>Examinations Institute</i>	A086
Avalia avaliações multimodais usando diferentes tipos de itens/questões abertas	A090
Avalia se a avaliação formativa é um preditor da avaliação somativa	A092
Avalia uma estratégia de avaliação baseado em competências	A113
Avalia exercícios criativos (ECs), uma técnica de avaliação aberta	A118
Avalia a implementação da avaliação diagnóstica	A124
Avalia os efeitos das avaliações formativas no desempenho dos alunos.	A039
Avalia como o fraseamento do enunciado pode recrutar compreensões dos alunos sobre as relações entre estrutura e propriedades	A134
Avalia a motivação, função, uso de instrumentos e desafios associados aos esforços de avaliação	A146

Avalia uma avaliação formativa, chamada <i>Who Wants To Be A Millionaire</i> , que possui o format game show	A195
Avalia itens bem projetados e testados de pesquisas publicadas anteriormente e repositórios de itens. A avaliação de mudança química e física (CPCA) foi projetada.	A058
Avalia a implementação de um método relativamente novo de pontuação e feedback automatizados em sala de aula	A042
Avalia a avaliação da competência dos alunos em argumentação científica escrita	A145

Fonte: os autores

Ao analisar o Quadro 22, nota-se a repetição do termo avaliação formativa em três artigos, no A092, A039 e A195. Os três artigos mencionam as avaliações formativas e seus efeitos nos alunos.

Trazemos abaixo um excerto do A039, para exemplificar o uso da avaliação formativa em questão e o contexto no qual a pesquisa foi realizada:

O objetivo deste estudo foi fornecer insights sobre os efeitos das avaliações formativas no desempenho durante um curso de química baseado em contexto sobre ácido láctico. Em um ambiente de pesquisa de ação em sala de aula, um projeto de grupo de controle pré-teste/pós-teste com replicações alternadas foi aplicado. O desempenho dos alunos foi medido em dois pré-testes, dois pós-testes e um teste de retenção. Os participantes eram alunos do 9º ano de uma escola secundária na Holanda. A análise de medidas repetidas mostrou um efeito significativo das avaliações formativas no desempenho dos alunos. Durante a implementação das avaliações formativas, discussões intrigantes surgiram entre os alunos, entre alunos e professores e entre professores. Adicionar avaliações formativas a abordagens baseadas em contexto reforça sua força para enfrentar os desafios atuais da educação em química. As avaliações formativas afetam o desempenho dos alunos positivamente e estimulam o feedback entre alunos e professores (A039, p. 155, tradução nossa)⁴⁸.

Observamos que o A039 avaliou os efeitos de avaliações formativas no desempenho dos alunos. Notou-se um efeito significativo das avaliações

⁴⁸ No texto original: "The aim of this study was to provide insights into the effects of formative assessments on achievement during a context-based chemistry course on lactic acid. In a classroom action research setting, a pre-test/post-test control group design with switching replications was applied. Student achievement was measured in two pre-tests, two post-tests and a retention test. Participants were Grade 9 students from one secondary school in the Netherlands. Repeated-measures analysis showed a significant effect of formative assessments on students' achievement. During the implementation of the formative assessments, intriguing discussions emerged between students, between students and teacher, and between teachers. Adding formative assessments to context-based approaches reinforces their strength to meet with the current challenges of chemistry education. Formative assessments affect students' achievement positively and stimulate feedback between students and teacher(s)" (A039, p. 155).

formativas no desempenho dos alunos, já que os estudantes se manifestaram por meio de discussões, justamente devido à natureza dialógica das avaliações formativas e a possibilidade de interação. Esses resultados corroboram resultados observado por Black e Wiliam (1998), ao terem observados impactos positivos na motivação de alunos e seus desempenhos, após uma análise de 580 artigos a respeito de avaliações formativas, especificamente o impacto de fornecer feedback direto aos alunos acerca do seu trabalho.

A importância do feedback na avaliação formativa para o impacto do desempenho também é comentada pelo A039:

[...] a eficácia das avaliações formativas depende fortemente de sua implementação na sala de aula, com um papel crucial do feedback (Filsecker e Kerres 2012). Durante as avaliações formativas descritas neste estudo, os alunos foram encorajados a fornecer feedback uns aos outros. O feedback do professor para os alunos foi fornecido imediatamente. Em geral, o feedback foi fornecido aos grupos e, se necessário, aos alunos individuais, e foi focado tanto na compreensão dos alunos do assunto (nível cognitivo) quanto em suas estratégias de aprendizagem (nível metacognitivo) (A039, p. 161-162, tradução nossa)⁴⁹.

Dessa forma, os resultados do A039 mostram os impactos positivos da utilização da avaliação formativa na Educação Básica, apresentando efeitos estatisticamente significativos no desempenho dos alunos e na aprendizagem, sugerindo conexões entre os impactos positivos e o feedback, seja ele realizado pelo professor para o aluno ou dos alunos entre si.

Analisando o Quadro 22, também se nota a repetição do termo “exercícios criativos”, como uma técnica de avaliação aberta, em dois artigos: o A077 e A118. O conceito de exercícios criativos não apareceu apenas na categoria C02, mas esteve presente em outras categorias do *corpus* também.

Incluímos abaixo um excerto do A118, para explicitar o conceito de exercícios criativos e comentar a respeito do contexto da pesquisa:

⁴⁹ No texto original: “[...] the efficacy of formative assessments strongly depends on their implementation in the classroom, with a crucial role of feedback (Filsecker and Kerres 2012). During the formative assessments described in this study, students were encouraged to provide feedback to each other. Feedback from teacher to students was provided immediately. In general, feedback was provided to groups, and if necessary to individual students, and was focused on both students’ understanding of the subject matter (cognitive level) and their learning strategies (metacognitive level)” (A039, p. 161-162).

Para promover a aprendizagem significativa, práticas de avaliação que incentivem a vinculação de conceitos precisam ser desenvolvidas e utilizadas. Exercícios Criativos (ECs) têm o potencial de encorajar tais vínculos. ECs são uma técnica de avaliação aberta onde os alunos recebem um único prompt e são solicitados a descrever o máximo de afirmações que puderem que sejam distintas, corretas e relevantes para o prompt. Este estudo descreve uma investigação qualitativa sobre as respostas dos alunos aos ECs para evidências de alunos vinculando conceitos ao longo do curso e a natureza dos conceitos vinculados. As descobertas indicam consideráveis interconexões de conteúdo nas respostas dos alunos. Além disso, os esforços dos alunos para fazer conexões revelaram vários equívocos sobre sua compreensão dos limites dos modelos. ECs são, portanto, propostos como um meio de encorajar os alunos a vincular conceitos e informar os instrutores sobre os links feitos, tanto corretamente quanto incorretamente (A118, p. 576, tradução nossa)⁵⁰.

Observamos que o foco do A118 é investigar o uso de exercícios criativos, uma técnica de avaliação aberta. O olhar dos autores foi direcionado às conexões ou vínculos criados entre conceitos pelos alunos, ao participarem de exercícios criativos.

Explorando os exercícios criativos, estes são um método de avaliação aberta e flexível sem um conjunto único ou limitado de respostas corretas. No geral, o exercício criativo fornece aos alunos um prompt relacionado ao tema estudado em questão, como por exemplo "um milhão de moléculas de SO_2 ", e pede que os alunos forneçam o máximo de declarações distintas, precisas e relevantes possíveis sobre o prompt. Os alunos podem receber pontuações ou créditos para cada declaração válida que fizerem. A avaliação pode ser estruturada por meio do professor, com o estabelecimento de um número especificado de declarações necessárias para alcançar o crédito/pontuação total, simultaneamente definindo um limite no crédito/pontuação total possível para a tarefa.

Um benefício importante dos exercícios criativos é que eles motivam os alunos a conectarem conceitos aprendidos anteriormente com conceitos ou

⁵⁰ No texto original: "To promote meaningful learning, assessment practices that encourage the linking of concepts need to be developed and utilized. Creative Exercises (CEs) have the potential to encourage such links. CEs are an open-ended assessment technique where students are given a single prompt and are asked to describe as many statements as they can that are distinct, correct, and relevant to the prompt. This study describes a qualitative investigation into student responses to CEs for evidence of students linking concepts throughout the course and the nature of the linked concepts. The findings indicate considerable interconnections of content in student responses. Further, students' efforts toward making connections revealed several misconceptions regarding their understanding of the limits of models. CEs are therefore proposed as a means to encourage students to link concepts and to inform instructors about the links made, both correctly and incorrectly" (A119, p. 576).

contextos novos. De acordo com o A118, analisar as respostas dos alunos aos exercícios criativos revela uma gama de conceitos vinculados em Química Geral, destacando não só as conexões efetivas, mas também as inadequadas. Portanto, exercícios criativos como um método avaliativo pode ajudar os professores a compreenderem as ligações conceituais dos alunos e enfatizarem a importância de fazer essas conexões. As respostas incorretas dos alunos também podem desencadear discussões em sala de aula sobre as limitações dos modelos aprendidos e ajudar a desenvolver novos prompts para outros exercícios criativos.

Como mencionado anteriormente, os exercícios criativos figuraram em outras categorias do *corpus* também, se mostrando como um conceito de interesse dos pesquisadores em Avaliação na área de Ensino de Química, assim como o conceito de avaliações formativas, por exemplo.

Ainda na categoria C02, destacamos o A026, por expandir os resultados da *AmERICan Chemical Society's Examinations Institute* (ACS Exams). A pesquisa do instituto em questão é considerada importante, devido ao seu alcance e impacto no contexto norte-americano para o tema em Avaliação na área de Ensino de Química.

Abaixo apresentamos um excerto do A026 para discutir seu contexto:

A avaliação continua a desempenhar um papel crescente no ensino superior. [...] Considerando a mudança do papel do professor na avaliação, o *AmERICan Chemical Society's Examinations Institute* (ACS Exams) renova seu interesse em entender como as avaliações estão sendo desenvolvidas e usadas em experiências instrucionais. Em 2011, o ACS Exams relatou no *Assessment Update* um "retrato" dos esforços atuais de avaliação na educação química pós-secundária (Emenike et al. 2011). Agora, seis anos após esse relatório, revisamos nosso estudo de pesquisa com ênfase na exploração das escolhas que o corpo docente de química faz ao usar um conjunto de avaliações em um curso no qual eles têm mais influência. É neste ponto que um novo retrato pode ajudar os educadores químicos e aqueles na indústria química a apoiar instituições de ensino superior no interesse de avaliar o aprendizado dos alunos. Nosso trabalho explora os diferentes tipos de avaliações em cursos de nível inferior e superior, bem como em cursos de matrícula pequena e grande (A026, p. 1, tradução nossa)⁵¹.

⁵¹ No texto original: "Assessment continues to play an expanding role in higher education. Faculty, more than ever before, are expected to be personally involved in generating and evaluating assessment data for internal and external reviews, including for program and institutional accreditation. Considering the changing role of the instructor in assessment, the American Chemical Society's Examinations Institute (ACS Exams) renews its interest in understanding how assessments are being developed and used across instructional experiences. In 2011, ACS Exams reported in *Assessment Update* a "snapshot" of

Assim, o A026 avaliou o sistema avaliativo de diferentes contextos norte-americanos, relatados por 1282 docentes de Química e/ou Ensino de Química em 2016. Os participantes responderam a 28 itens a respeito de suas práticas de ensino, suas crenças sobre ensino e aprendizagem, o contexto de sua instituição e de sua sala de aula. Especificamente, os entrevistados descreveram os cursos de química a qual pertenciam, que ministraram nos últimos três anos, considerando a data da pesquisa. Com isso o foco do A026 esteve em identificar: "Quais ferramentas de avaliação foram utilizadas nos cursos identificados?"

O interesse do A026 reside em identificar como as avaliações têm sido desenvolvidas em sala de aula por meio da análise das escolhas feitas por professores em suas práticas avaliativas. O A026 explora os diferentes tipos de avaliações em diferentes contextos, como: cursos de nível inferior e superior, bem como cursos com um elevado e baixo número de matriculados.

Conforme o A026, a ferramenta de avaliação mais usada e descrita pelos entrevistados foi "avaliações escritas pessoalmente", seguida por "exames finais escritos pessoalmente" e "questionários escritos pessoalmente". No total, 15 ferramentas foram mencionadas. Em ordem decrescente, estas foram: avaliações escritas pessoalmente; exames finais escritos pessoalmente; questionários escritos pessoalmente; tarefas de casa escritas; tarefas de casa online; o exame final do ACS; relatórios escritos e trabalhos de conclusão de curso; sistemas de resposta do aluno; apresentações dos alunos; pesquisas (por exemplo, de autoeficácia); exames comuns; exame finais comuns; questionários comuns; mapas conceituais; e testes ou inventários de conceitos.

A respeito das variações entre os contextos, trabalhos de casa escritos, relatórios escritos e trabalhos de conclusão de curso, e apresentações de

current assessment efforts in postsecondary chemistry education (Emenike et al. 2011). Now six years since that report, we have revised our survey study with an emphasis on exploring the choices chemistry faculty make in using a suite of assessments in a course for which they have the most influence. It is at this point that a new snapshot can help chemical educators, and those in the Chemical industry, to support institutions of higher education in the interest of assessing student learning. Our work explores the differing types of assessments in lower- and upperlevel courses, as well as within small- and large-enrollment courses. Several themes emerge from these data, including the preponderance of online homework systems in large courses, written reports and term papers in small courses, and surveys (e.g., selfefficacy) in lower-level courses. The value and utility of assessments are indicated by their use in differing contexts" (A026, p. 1).

alunos foram mais frequentemente usados em cursos com menor número de alunos. Já os cursos com maior número de alunos incorporaram trabalhos de casa online.

Avaliações coordenadas, ou seja, avaliações compostas por questionários ou exames aplicados por todos os instrutores que ensinam em um curso na instituição, foram excessivamente usadas nos cursos, e em combinação com sistemas de resposta de alunos, como “clickers”⁵².

Assim, os resultados do A026 fornecem um retrato das ferramentas de avaliações mais utilizadas pelos participantes da pesquisa e mostram que o valor e a utilidade das avaliações são indicados por seu uso em diferentes contextos.

4.2.3 Avalia propostas de ensino (C03)

A categoria C03 consistiu dos artigos que avaliaram propostas de ensino. No Quadro 23 apresentamos todas as subcategorias encontradas na categoria C03 e os respectivos códigos desses artigos. As subcategorias mostram as propostas de ensinamentos específicas que foram avaliadas.

Quadro 23 - C03: Avalia propostas de ensino⁵³

Subcategoria	Código
Avalia soluções simuladas de alunos versus soluções reais na avaliação por pares	A004
Avalia uma tarefa/intervenção ácido-base de <i>Writing to Learn</i>	A016
Avalia frameworks de argumentação teóricos para melhorar a compreensão do aluno sobre interações não covalentes	A020
Avalia uma intervenção pedagógica baseada em <i>storytelling</i>	A028
Avalia abordagens de aprendizado do aluno, que podem prever suas chances de sucesso no curso	A030
Avalia e investiga como as orientações de ensino de ciências dos professores de química em formação mudam durante uma intervenção de dois semestres projetada para aprimorar seu conhecimento pedagógico do conteúdo para o ensino da natureza da ciência	A040
Avalia e examina a eficácia da aprendizagem colaborativa com dicas e a tutoria entre colegas com dicas e a aprendizagem individual com dicas	A041

⁵² Clickers, também conhecidos como sistemas de resposta do público ou sistemas de resposta do aluno, são dispositivos portáteis ou aplicativos de software usados no Ensino de Ciências para facilitar o aprendizado interativo. Os instrutores fazem perguntas durante as aulas, relacionadas a conceitos-chave ou tópicos sendo discutidos e os alunos usam seus clickers para enviar respostas anonimamente. Isso pode ser feito usando dispositivos dedicados ou aplicativos de smartphone. As respostas são coletadas e exibidas em tempo real, permitindo que os professores avaliem a compreensão e ajustem seu ensino (Sevian; Robinson, 2011).

⁵³ Consideramos como proposta de ensino; abordagem, tarefa, intervenção, framework e workshop.

Avalia a aprendizagem cooperativa ⁵⁴	A046
Avalia um curso de experiência escolar baseado em conhecimento pedagógico do conteúdo	A062
Avalia uma intervenção composta de exercícios criativos	A067
Avalia workshops	A069
Avalia a abordagem construtivista variada	A071
Avalia duas intervenções: uma lista de verificação de autoavaliação e um vídeo gravado pelo aluno	A073
Avalia a intervenção em relação ao Esquema Perry de Desenvolvimento Intelectual	A082
Avalia o efeito de uma estrutura de curso de aprendizado ativo	A094
Avalia uma abordagem de investigação científica baseada em argumentação	A097
Avalia o impacto da aprendizagem cooperativa (uma intervenção pedagógica)	A102
Avalia questões baseadas na Teoria da Carga Cognitiva como uma abordagem	A103
Avalia uma abordagem de aprendizagem combinada	A105
Avalia a aprendizagem aprimorada por teste	A106
Avalia o impacto de um exercício prático de laboratório que usou kits de modelos moleculares	A109
Avalia um novo módulo para alunos do ensino médio para o ensino de química quântica	A116
Avalia práticas autênticas de modelagem química e contextos para uso no ensino de química	A149
Avalia um projeto de pesquisa-ação de professores em que uma abordagem de <i>Systems Thinking</i> foi usada	A150
Avalia disciplinas com um modelo <i>Flipped Classroom</i>	A152
Avalia o efeito de ensinar química durante todo o ano letivo do ensino médio utilizando o raciocínio abstrato	A163
Avalia o modelo de sala de aula invertida	A164
Avalia uma abordagem de métodos mistos de mapas conceituais e exercícios criativos	A165
Avalia as abordagens de aprendizagem em graduandos de química.	A169
Avalia o momento de um experimento no programa de laboratório	A172
Avalia até que ponto o tipo de instrução (abordagem) usada durante um curso de laboratório de química geral afeta a capacidade dos alunos de usar ideias centrais para se envolver em práticas científicas	A187
Avalia se ao criar um vídeo de técnicas laboratoriais os alunos apresentam uma maior retenção de informações, em um curso de laboratório	A191
Avalia as unidades instrucionais que integram modelagem molecular computacional e visualização para ensinar conceitos químicos fundamentais	A192
Avalia como o Processo de Hierarquia Analítica resulta em um ensino de Química mais eficaz	A081

⁵⁴ Este artigo focou nos efeitos da aprendizagem cooperativa. Comparou o desenvolvimento do conhecimento do conteúdo de química e habilidades transferíveis ou genéricas entre os alunos em: um formato baseado em palestra versus em um formato de aprendizagem cooperativa.

Avalia o efeito da implementação da estratégia “Abordagem Sistêmica de Ensino e Aprendizagem (SATL)”	A202
Avalia uma tarefa de construção de cadeia de raciocínio	A205
Avalia o valor do contato direto e execução (enactment) ao aprender química usando modelos moleculares 3D em vídeo e aulas em sala de aula	A072
Avalia o efeito da modalidade instrucional inicial e do conhecimento prévio na aprendizagem	A100
Avalia a introdução de como usar cartas de controle com um experimento simples de titulação durante o Laboratório de Análise Quantitativa	A184

Fonte: os autores

Observamos que houve um significativo número de artigos que avaliaram propostas de ensino. De fato, a categoria C03 foi a que concentrou o maior número de artigos (39 artigos), representando 19% do corpus. Portanto, existe uma tendência nas pesquisas em Avaliação na área de Ensino de Química de avaliar propostas de ensino. A estrutura dos artigos que estão na categoria C03 é motivada pelo interesse dos pesquisadores em elaborar e/ou aplicar uma proposta de ensino específica em um contexto educacional particular e, posteriormente avaliar seu impacto nos alunos, sua qualidade, sua construção, suas características, seus efeitos, etc.

Vale ressaltar outro ponto que pode ter favorecido a grande incidência de artigos nessa categoria. O termo proposta de ensino é consideravelmente abrangente, englobando uma grande possibilidade de mediação em sala de aula como: uma abordagem, uma tarefa, uma intervenção, um framework, um workshop, uma atividade, uma sequência didática, etc. Também, ao analisar o Quadro 23, é evidente a grande variedade de propostas de ensino avaliadas. Dessa forma, a abrangência do termo proposta de ensino, a pluralidade de propostas evidenciadas e o próprio interesse dos pesquisadores em conduzir esse tipo de estudo podem ter contribuído para explicar esses fenômenos observados no *corpus*.

Ao analisar o Quadro 23, notamos a repetição de três termos em mais de um artigo. A aprendizagem cooperativa nos artigos A046 e A102; a sala de aula invertida ou *flipped classroom* nos artigos A152 e A164; e os exercícios criativos nos artigos A067 e A165.

Trazemos abaixo um excerto do A016 para exemplificar os artigos alocados a categoria C03.

Química ácido-base é um tópico fundamental que é ensinado em cursos em todo o currículo de química. Os alunos geralmente têm dificuldade em distinguir entre as diferentes teorias de química ácido-base — química ácido-base de Brønsted–Lowry e Lewis — e aplicar essas duas definições corretamente em cenários desconhecidos. Para ajudar os alunos a aprender essas definições e serem capazes de aplicá-las, uma tarefa de *Escrita para Aprender* sobre ácido-base foi desenvolvida e avaliada (A016, p. 383, tradução nossa, grifo nosso)⁵⁵.

Nota-se que o A016 avaliou uma proposta de ensino, a *Writing to Learn* (WTL), ou *Escrita para Aprender*, no português. O WTL é uma abordagem pedagógica que promove uma compreensão mais profunda de assuntos por meio da escrita. De acordo com o A016 as tarefas WTL podem aumentar significativamente a compreensão dos alunos sobre conceitos específicos de química, pois exigem dos alunos uma maior concentração para identificar e relatar suas explicações em vez de apenas escrever mecanicamente sobre um tema.

No A016, alunos matriculados em uma disciplina de Laboratório em Química Orgânica se envolveram em uma atividade WTL onde definiram e conectaram dois modelos teórico ácido-base, com a intenção de permitir aos alunos a aplicação de conceitos fundamentais em um novo contexto. Além disso, o processo WTL exigiu dos alunos uma revisão por pares. Isso permitiu que os mesmos fornecessem e recebessem feedback. De acordo com o A016 essa inclusão demonstrou melhorar a qualidade e o conteúdo da escrita dos alunos.

O objetivo do A016 foi avaliar como a participação na atividade WTL influenciou a compreensão conceitual e a confiança dos alunos em abordar problemas que envolvem o conceito ácido-base. Para isso comparou-se um grupo de tratamento, que concluiu a tarefa WTL, a um grupo de comparação que realizou uma tarefa diferente.

A tarefa WTL focou na contraindicação da levotiroxina⁵⁶, permitindo que os alunos definissem e aplicassem as teorias ácido-base de Brønsted–Lowry e Lewis em um contexto real. Os alunos tiveram que escrever um e-mail como se fossem químicos medicinais conversando com um médico a respeito do medicamento.

⁵⁵ No texto original: “Acid–base chemistry is a foundational topic that is taught in courses across the chemistry curriculum. Students often have difficulty distinguishing between the different theories of acid–base chemistry—Brønsted–Lowry and Lewis acid–base chemistry—and applying these two definitions correctly in unfamiliar scenarios. To help students learn these definitions and be able to apply them, an acid–base Writing-to-Learn assignment was developed and evaluated” (A016, p. 383).

⁵⁶ A levotiroxina é um hormônio sintético da tireoide, empregado principalmente no tratamento do hipotireoidismo, uma condição em que a glândula tireoide não gera hormônios suficientes.

Assim a tarefa WTL promoveu/simulou um diálogo científico autêntico que poderia ocorrer no mundo real. A tarefa foi parte de uma série de atividades que buscou aprofundar a compreensão conceitual dos alunos por meio da escrita, incorporando três componentes principais: um rascunho inicial, a revisão por pares e um rascunho revisado. Por outro lado, o grupo de comparação concluiu planilhas mais curtas com menos profundidade, incentivando a discussão em grupo, mas sem o contexto simulado da tarefa WTL.

A respeito da coleta de dados, essa etapa envolveu: perguntas pré e pós-conceituais, a própria escrita dos alunos e entrevistas reflexivas para analisar as compreensões dos alunos. A análise incluiu métodos quantitativos e qualitativos, gerando 186 respostas do grupo WTL e 335 do grupo de comparação.

A respeito dos resultados, as análises quantitativas mostraram que a participação no WTL foi capaz de prever, significativamente, um desempenho melhor em explicações pós-questionário. Os alunos que participaram do WTL mostraram um aumento médio de 0,149 níveis de rubrica. Além disso, percebeu-se que os alunos que tinham participado do WTL tiveram maior probabilidade de responder corretamente às perguntas sobre acidez de Lewis.

Também foi evidenciado melhorias estatísticas em todas as categorias de rubrica, particularmente nos conceitos de Brønsted–Lowry dentre os alunos participantes do WTL. Outro resultado foi a importância do processo de revisão por pares, já que este facilitou a compreensão dos alunos dos conceitos ácido-base de Lewis. Além disso, os resultados mostraram que a participação no WTL aumentou significativamente as pontuações de confiança dos alunos. Desse modo, o A016 conclui que os resultados mostram que tarefas WTL podem melhorar a compreensão conceitual e a confiança em tópicos desafiadores de química orgânica, caracterizando a proposta de ensino como uma opção viável para uso em sala de aula.

Trazemos abaixo um excerto do A028 para continuarmos explicitando os artigos alocados a categoria C03.

Uma abordagem de storytelling tem sido vista como uma forma poderosa de ensinar ciências e despertar interesse e promover atitudes positivas em relação à aprendizagem de ciências nos primeiros anos. O objetivo do nosso estudo foi determinar como os alunos do Key Stage 3 (KS3)⁵⁷ com idades entre 12 e 14 anos em

⁵⁷ KS3 são alunos entre 12-14 anos de idade. Em Portugal, a educação formal em química começa neste momento – a partir do Key Stage 3 (KS3) (A028).

escolas portuguesas – vivenciaram a aprendizagem de química por meio de narrativas e como eles, por sua vez, vivenciaram a criação de histórias usando uma abordagem de narrativa com crianças pré-escolares. Nosso objetivo foi perceber a apropriação de conceitos de química pelas crianças pré-escolares por meio de seus desenhos, os resultados coletados durante a intervenção pedagógica e as gravações das discussões entre as crianças pré-escolares, os alunos e os professores da pré-escola. Os alunos do KS3 também receberam um questionário de autoavaliação como forma de avaliar a dinâmica pedagógica e a motivação para aprender química (A028, p. 302, tradução nossa)⁵⁸.

Observa-se que o A028 busca analisar e avaliar uma proposta de ensino específica, o uso de *Storytelling*. De acordo com o A028 o *Storytelling* é caracterizado como um método envolvente que aprimora a educação científica, sendo inicialmente empregado no Ensino de Química por Wally *et al.* (2005), onde fez-se uma conexão entre a popular série "Harry Potter" com aulas de química para alunos do ensino fundamental. No estudo supracitado, combinou-se experimentos práticos com narrativas familiares da série para promover o entusiasmo dos alunos.

O A028 buscou envolver alunos do Ensino Fundamental, que estavam aprendendo sobre conceitos ácido-base, na criação de histórias e atividades práticas para alunos pré-escolares, aumentando a compreensão a respeito do tema para ambos os grupos. Nesse sentido, o A028 buscou explorar: quais histórias e atividades os alunos do Ensino Fundamental desenvolveram para os alunos pré-escolares; as experiências no Ensino de Química que os alunos do Ensino Fundamental tiveram por meio de contação de histórias; e os conceitos de química que os alunos pré-escolares demonstraram após se envolverem na atividade.

A respeito do método da pesquisa, os participantes consistiram de 53 alunos do Ensino Fundamental (KS3) (idades de 13 a 14) e 37 alunos pré-escolares (idades de 4 a 5). As interações ocorreram em seis aulas de química e duas sessões de *storytelling* em três etapas. As três etapas foram: Escrita da história; Interação com os pré-escolares; e Feedback. No contexto da pesquisa isso foi realizado da seguinte

⁵⁸ No texto original: "A storytelling approach has been seen as a powerful way to teach science and arouse interest and promote positive attitudes toward learning science in the early years. The purpose of our study was to determine how middle school students – Key Stage 3 (KS3) aged 12–14 in Portuguese schools – experienced learning chemistry through storytelling and how they, in turn, experienced creating stories using a storytelling approach with pre-school children. We aimed to perceive the appropriation of concepts of chemistry by the pre-school children through their drawings, the results collected during the pedagogical intervention and the recordings of the discussions between the pre-school children, the students and the pre-school teachers. The KS3 students were also given a self-assessment questionnaire as a way of assessing the pedagogical dynamics and the drive and motivation to learn chemistry" (A028, p. 302).

maneira: primeiro, o professor criou uma história e atividades para ensinar conceitos ácido-base para os alunos do Ensino Fundamental. Essa história serviu como um modelo para os alunos do Ensino Fundamental. Em seguida, os alunos do Ensino Fundamental desenvolveram suas próprias histórias e atividades para os alunos mais novos, os pré-escolares. Após o momento de *storytelling* entre os alunos mais velhos e mais novos, os últimos ilustraram, por meio de desenhos, seus entendimentos que foram posteriormente analisados.

O impacto da proposta de ensino foi avaliado em três estágios: introdução do conceito, preparação do material e interações com pré-escolares. Para isso utilizou-se estatísticas descritivas e avaliações qualitativas das histórias contadas pelos alunos do Ensino Fundamental e os desenhos realizados pelos alunos pré-escolares após os momentos de *storytelling*.

A respeito dos resultados, estes alinharam-se com a literatura existente sobre *storytelling*, indicando que contar histórias e realizar atividades práticas motivam efetivamente as crianças a aprender ciências, particularmente química. Os alunos do Ensino Fundamental responderam positivamente à história utilizada e suas atividades relacionadas, sugerindo um impacto positivo em seu aprendizado de química. Além disso, percebeu-se como a proposta motivou os alunos do Ensino Fundamental a criarem seus próprios materiais educacionais.

As histórias criadas pelos alunos abordaram conceitos científicos como estados da matéria e mudanças climáticas de maneira efetiva, adaptando-os para manter a acessibilidade para os alunos mais jovens. Apesar disso, o A028 destaca escolhas não convencionais nas estruturas narrativas das histórias dos alunos do Ensino Fundamental, o que afetou o engajamento em determinados momentos.

A interação entre alunos do Ensino Fundamental e os alunos pré-escolares foi apontado como um ponto alto da experiência, já que os alunos mais velhos tiveram sucesso ao realizarem suas pesquisas e apresentarem os conceitos com uma linguagem mais simples, enquanto as crianças se engajaram na atividade.

A análise dos desenhos dos alunos pré-escolares mostrou que o conhecimento científico foi efetivamente compreendido. Para isso o A028 indica a presença de elementos científicos nos desenhos dos alunos que sugerem que os alunos mais velhos comunicaram com sucesso os principais conceitos científicos.

Em resumo, o A028 relata que o projeto demonstrou as potencialidades na combinação de *storytelling* e atividades práticas e foi eficaz para ensinar química tanto para alunos do Ensino Fundamental quanto para alunos pré-escolares. A interação facilitou a aprendizagem dos conhecimentos e promoveu um ambiente envolvente entre os alunos. Para futuras implementações, recomenda-se uma atenção redobrada no momento de ajustar as atividades às necessidades do público infantil, assegurando uma adaptação da comunicação científica para suas necessidades específicas da idade.

4.2.4 Avalia um instrumento, ferramenta ou tecnologia (C04)

A categoria C04 consistiu dos artigos que avaliaram um instrumento, ferramenta ou tecnologia. No Quadro 24 apresentamos todas as subcategorias encontradas na categoria C04 e os respectivos códigos desses artigos. As subcategorias mostram os instrumentos, ferramentas e tecnologias que foram avaliadas.

Quadro 24 - C04: Avalia um instrumento, ferramenta ou tecnologia⁵⁹

Subcategoria	Código
Avalia um instrumento ou ferramenta	
Avalia mapas conceituais como ferramentas para avaliar	A022, A074, A189
Avalia o instrumento Avaliação de Desempenho de Experiências de Pesquisa de Graduação (PURE)	A054
Avalia uma ferramenta para mensurar o pensamento crítico em química	A059
Avalia uma ferramenta para avaliar as respostas escritas dos alunos no conteúdo de ácido-base de Lewis	A060
Avalia a avaliação do Nível de proficiência Celular (GCPA)	A061
Avalia ferramentas de diagnóstico	A083
Avalia a validade do instrumento Test of Science Related Attitudes (TOSRA)	A089
Avalia o Math-Up Skills Test (MUST) como uma ferramenta potencial para identificar alunos em risco	A101
Avalia o instrumento de Avaliação dos Ganhos de Aprendizagem do Aluno (SALG)	A115
Avalia diálogos cogenerativos como uma ferramenta avaliativa	A123

⁵⁹ Todos os instrumentos ou ferramentas que envolvem mídia, foram deslocados para a subcategoria de tecnologias.

Avalia a eficácia do ACT Mathematics Test e o Whimbey Analytical Skills Inventory (WASI)	A133
Avalia se a ordem das perguntas estavam presentes em um inventário de conceitos sobre química à base de ácido tem algum efeito	A139
Avalia o <i>Thermochemistry Concept Inventory</i> (TCI)	A141
Avalia um instrumento para avaliar o uso da metacognição	A144
Avalia uma tarefa com um mapa conceitual	A148
Avalia as respostas dos alunos ao <i>immediate feedback assessment technique</i> (IF-ATs)	A153
Avalia a confiabilidade de ítems no <i>Adaptive Chemistry Assessment Survey for Teachers</i>	A158
Avalia a versão turca do <i>Anxiety Rating Scale</i> (DCARS)	A159
Avalia como um instrumento de avaliação com uma estrutura única pode ser usado para identificar ideias incorretas prévias	A175
Avalia o framework <i>Representation Mapping</i>	A180
Avalia e constrói <i>construct map framework</i> , que orienta o desenvolvimento de recursos instrucionais e de avaliação relacionados à base física de interações não covalentes em um contexto bioquímico	A057
Avalia uma tecnologia	
Avalia a tecnologia de aplicação de interfaces naturais de usuário (sensor Kinect) na criação de um laboratório químico virtual	A027
Avalia um ambiente virtual baseado em um laboratório de química como ferramenta	A068
Avalia um software para investigar o conhecimento estrutural dos alunos	A078
Avalia tecnologias adaptativas para aumentar a participação prática de alunos com cegueira ou baixa visão	A087
Avalia o impacto de uma simulação computacional colaborativa de pH	A107
Avalia como três tecnologias foram utilizadas em combinação para alinhar o aprendizado e a avaliação do aluno	A117
Avalia um ambiente de aprendizagem baseado em computador	A126
Avalia modelos de computador como avaliação formativa na compreensão do aluno sobre a natureza dos modelos	A162
Avalia o ChemVLab +, um conjunto de atividades de química online	A194
Avalia a usabilidade de um instrumento de diagnóstico baseado na web	A006
Avalia o Kahoot como uma ferramenta	A179
Avalia um ambiente de resolução de problemas baseado em computador	A201
Avalia uma interface baseada em cubo na tabela periódica virtual	A049
Avalia um curso híbrido de Simetria Molecular e Teoria de Grupos que combina instrução presencial tradicional com um ambiente de aprendizagem on-line aprimorado pela web	A051

Fonte: os autores

Nota-se que os artigos alocados na categoria C04 avaliaram ferramentas, instrumentos ou tecnologias. A categoria C04 foi a segunda mais representativa do *corpus*, englobando 36 artigos, representando 17,5% do *corpus*.

Assim, há uma tendência nas pesquisas em Avaliação na área de Ensino de Química em avaliar instrumentos, ferramentas ou tecnologias.

Os instrumentos e ferramentas avaliados foram: mapas conceituais, PURE⁶⁰, uma ferramenta para mensurar o raciocínio crítico; uma ferramenta para avaliar respostas escritas; GCPA⁶¹; ferramentas de diagnóstico; TOSRA⁶²; MUST⁶³; SALG⁶⁴; diálogos cogenerativos; o ACT Mathematics Test⁶⁵; WASI⁶⁶; um inventário de conceitos; TCI⁶⁷; um instrumento para avaliar o uso da metacognição; uma ferramenta com um mapa conceitual; IF-ATs⁶⁸; o Adaptive Chemistry Assessment Survey for Teachers; DCARS⁶⁹; um instrumento de avaliação com uma estrutura única; Representation Mapping; e o construct map framework.

As tecnologias avaliadas foram: um sensor Kinect na criação de um laboratório químico virtual; um ambiente virtual baseado em um laboratório; um software para investigar o conhecimento estrutural; tecnologias adaptativas para alunos com cegueira ou baixa visão; uma simulação computacional colaborativa; a combinação de tecnologias para avaliação e aprendizagem; um ambiente de aprendizagem baseado em computador; modelos de computador como avaliação formativa; um conjunto de atividades de química online chamado ChemVLab +; um instrumento de diagnóstico baseado na web; o Kahoot; um ambiente de resolução de problemas baseado em computador; uma interface baseada em cubo; e um ambiente de aprendizagem on-line aprimorado pela web.

Ao analisar o Quadro 24, especificamente os artigos que avaliaram instrumentos e ferramentas, observamos a repetição do termo mapa conceitual em 4 artigos: A022, A074, A189 e A148. Esses artigos avaliaram o uso de mapas

⁶⁰ O instrumento “Avaliação de Desempenho de Experiências de Pesquisa de Graduação”.

⁶¹ A ferramenta “Nível de proficiência Celular”.

⁶² O instrumento “Teste de atitudes relacionadas à Ciência”.

⁶³ A ferramenta “Math-Up Skills Test”, utilizada para identificar alunos em risco.

⁶⁴ O instrumento “Avaliação dos Ganhos de Aprendizagem do Aluno”

⁶⁵ O Teste de Matemática ACT é uma avaliação padronizada que faz parte do exame ACT (American College Testing), frequentemente usado para admissões em faculdades nos Estados Unidos.

⁶⁶ O Whimby Analytical Skills Inventory (WASI) é uma ferramenta para avaliar o pensamento analítico e as habilidades de resolução de problemas.

⁶⁷ Uma ferramenta de avaliação educacional projetada para avaliar a compreensão dos alunos sobre conceitos-chave em termoquímica. Ela visa identificar equívocos comuns e avaliar a profundidade do conhecimento dos alunos.

⁶⁸ O “Immediate Feedback Assessment Technique” (IF-AT) é uma ferramenta educacional interativa para aumentar o envolvimento e a aprendizagem dos alunos por meio de feedback imediato durante as avaliações

⁶⁹ Uma ferramenta projetada para avaliar os níveis de ansiedade em indivíduos, particularmente em ambientes educacionais e clínicos.

conceituais como ferramentas avaliativas. A fim de exemplificar a natureza dos artigos da C04 e em especial a avaliação de mapas conceituais, devido a sua maior incidência dentro da categoria, apresentamos um excerto do A022:

Embora os mapas conceituais tenham sido usados como métodos alternativos de avaliação na educação, tem havido um debate em andamento sobre como avaliar os mapas conceituais dos alunos. Este estudo discute como avaliar os mapas conceituais dos alunos como uma ferramenta de avaliação antes e depois de 15 tarefas de Prever-Observar-Explicar (CB-POE) baseadas em computador relacionadas à química ácido-base. 12 alunos do ensino médio participaram deste estudo. Os mapas conceituais dos alunos foram avaliados levando em consideração três partes: nível qualitativo, quantitativo e representacional (A022, p. 843, tradução nossa)⁷⁰.

A pesquisa do A022 foi realizada com alunos do Ensino Médio da Turquia. Nela, os mapas conceituais foram utilizados para coletar dados antes e depois de tarefas específicas sobre o tema ácido-base. Os alunos criaram mapas conceituais pré-tarefas e pós-tarefas, com um intervalo de duas semanas entre cada etapa. No que diz respeito ao tipo de mapa conceitual utilizado, os autores comentam que optaram por mapas do tipo Construct-a-map para avaliar a compreensão conceitual dos alunos, devido a essa técnica ser mais sensível para capturar a estrutura do conhecimento dos alunos.

Os resultados do A022, no que diz respeito às análises qualitativas dos mapas conceituais, indicam que a maioria dos alunos construiu seus mapas conceituais pré e pós de forma não hierárquica; enquanto eles elaboraram mapas conceituais mais interconectados após as tarefas aplicadas no estudo. No que diz respeito às análises quantitativas, estas mostraram que houve uma diferença significativa entre as pontuações dos mapas conceituais pré e pós dos alunos ($z = 3,05$; $p < 0,05$). Em relação ao nível representacional utilizado pelos alunos, estes usaram o nível macroscópico com mais frequência do que os níveis microscópico e simbólico em seus mapas conceituais, tanto nos pré e pós testes. Apesar disso, os alunos aumentaram o número de representações de nível macroscópico, microscópico e simbólico em seus mapas conceituais posteriores.

⁷⁰ No texto original: "Although concept maps have been used as alternative assessment methods in education, there has been an ongoing debate on how to evaluate students' concept maps. This study discusses how to evaluate students' concept maps as an assessment tool before and after 15 computer-based Predict–Observe–Explain (CB-POE) tasks related to acid–base chemistry. 12 high school students participated in this study. Students' concept maps were evaluated taking into account three parts: qualitative, quantitative and representational level" (A022, p. 843).

Em relação a maneira que os mapas conceituais foram avaliados, o A022 discute que estes foram avaliados em três estágios: análise quantitativa, qualitativa e de nível representacional. Para a análise quantitativa, utilizou-se uma “análise de critérios”, composta de seis critérios: “conceitos, proposições, hierarquias, ligações cruzadas, exemplos e grau de interconexão”, baseando se parcialmente nas pesquisas de Novak e Gowin (1984), Markham *et al.* (1994) e Martin *et al.* (2000) para alguns dos critérios.

Outra característica avaliada nos mapas conceituais foram as proposições estabelecidas entre dois conceitos. Nesse sentido os autores estabeleceram cinco categorias: compreensão sólida, compreensão parcial, concepção errônea, compreensão errada e em branco. Para a elaboração dessas categorias, basearam-se em Abraham *et al.* (1994). A partir das estruturações, pontuações foram atribuídas aos mapas conceituais dos estudantes.

O último estágio da avaliação dos mapas conceituais inclui a análise qualitativa. Para isso, os mapas dos alunos foram classificados como hierárquicos, não hierárquicos ou em cadeia. Essa classificação permitiu a visualização do conhecimento estrutural que os alunos possuíam. Para o A022, se os alunos construíssem mapas altamente interconectados, como rede ou hierárquicos com um considerável número de links e ligações cruzadas, isso demonstrava uma maior proficiência dos alunos no tema. Por outro lado, se os alunos construíssem mapas conceituais de cadeia com estruturas simples, isso demonstrava menor proficiência.

A partir do exposto, o A022 realiza a avaliação de mapas conceituais de alunos e demonstra um método para avaliação destes, conforme procedimentos qualitativos, quantitativos e de níveis representacionais. Assim, os mapas conceituais se configuram como ferramentas utilizadas não apenas para investigar a compreensão conceitual dos alunos, mas também para avaliar, de maneira válida e confiável sua compreensão.

A fim de continuarmos a exemplificação dos artigos alocados na categoria C04, trazemos um excerto do A059 abaixo:

Habilidades de pensamento crítico raramente são explicitamente avaliadas em universidades. Avaliações comerciais⁷¹ de pensamento

⁷¹ Avaliações comerciais de pensamento crítico são testes projetados para avaliar a capacidade de um sujeito de analisar, avaliar e sintetizar informações para tomar decisões informadas. Essas

crítico, que geralmente são genéricas em contexto, estão disponíveis. No entanto, a literatura sugere que avaliações que usam um contexto relevante para os alunos refletem com mais precisão suas habilidades de pensamento crítico. Este artigo descreve o desenvolvimento e a avaliação de um teste de pensamento crítico de química (o Danczak–Overton–Thompson Chemistry Critical Thinking Test ou teste DOT), definido em um contexto de química e projetado para ser administrado a alunos de graduação em química em qualquer nível de estudo. O desenvolvimento e a avaliação ocorreram em três versões do teste DOT por meio de uma variedade de fases quantitativas e qualitativas de testes de confiabilidade e validade (A059, p. 62, tradução nossa)⁷².

A partir do exposto, nota-se que o A059 busca desenvolver e avaliar um teste para avaliar o pensamento crítico em química. Para o A059 o pensamento crítico pode ser definido por meio de perspectivas de diferentes áreas, como a filosofia, a psicologia cognitiva e a educação. Na área educacional, o pensamento crítico concentra-se em comportamentos observáveis como identificar problemas e avaliar evidências (Bailin, 2002; Dressel; Mayhew, 1954). Apesar de haver diferentes definições para o pensamento crítico, existem temas comuns entre essas definições, como por exemplo: a necessidade de habilidades metacognitivas e comportamentos observáveis, o julgamento, o ceticismo e a racionalidade (Glaser, 1984; Kuhn, 1999).

O A059 destaca a necessidade de um teste de pensamento crítico em química que seja válido, confiável e livre de conhecimento específico da disciplina e que seja adequado para todos os níveis de alunos de graduação. Nesse sentido o A059 apresenta o Danczak–Overton–Thompson Chemistry Critical Thinking Test (DOT), construído para avaliar as habilidades de pensamento crítico de alunos de química. Assim, o A059 busca: apresentar sua metodologia para a construção do DOT e avaliá-lo como uma ferramenta de mensuração do pensamento crítico de alunos da graduação.

A respeito do método utilizado para desenvolver o DOT, este ocorreu em cinco estágios: *definição operacional*; *criação do teste piloto*; *teste de confiabilidade e validade*; *teste DOT V2*; e *teste DOT V3*. Na definição operacional,

avaliações são frequentemente usadas em ambientes educacionais, ambientes corporativos e para desenvolvimento profissional (A059).

⁷² No texto original: “Critical thinking skills are seldom explicitly assessed at universities. Commercial critical thinking assessments, which are often generic in context, are available. However, literature suggests that assessments that use a context relevant to the students more accurately reflect their critical thinking skills. This paper describes the development and evaluation of a chemistry critical thinking test (the Danczak–Overton–Thompson Chemistry Critical Thinking Test or DOT test), set in a chemistry context, and designed to be administered to undergraduate chemistry students at any level of study. Development and evaluation occurred over three versions of the DOT test through a variety of quantitative and qualitative reliability and validity testing phases” (A059, p. 62).

buscou-se elaborar uma definição operacional de pensamento crítico por meio da comparação de dados dos participantes com a literatura existente. Neste momento, o intuito foi criar uma definição funcional que informasse o teste DOT. Na *criação do teste piloto*, elaborou-se uma versão piloto do teste DOT a partir dos insights do primeiro estágio. Durante o *teste de confiabilidade e validade* (DOT V1), a confiabilidade interna e a validade de conteúdo da primeira versão (DOT V1) foram avaliadas. Durante o *teste DOT V2* uma amostra de alunos de graduação em química realizou o DOT V2 e o Watson-Glaser Test of Critical Thinking Appraisal (WGTCAS) para avaliar a confiabilidade do teste-reteste, a validade convergente e a validade de conteúdo. Durante o *teste final* (DOT V3), a versão final (DOT V3) passou por extensas avaliações de confiabilidade e validade, administradas a vários grupos de alunos e pesquisadores.

A respeito dos resultados do A059, o desenvolvimento e avaliação do DOT consistiu de três iterações da ferramenta. No DOT V1 notou-se uma consistência interna limitada. Além disso, os grupos focais revelaram problemas com: clareza da instrução, o que se manifestou em interpretações erradas das instruções; redação das perguntas, devido a frases confusas; qualidade da informação, já que não havia contexto suficiente nas perguntas; e conhecimento prévio, relacionado a uma dependência excessiva de conhecimento externo.

Após os ajustes necessários, no DOT V2, observou-se boa confiabilidade do teste-reteste. Os participantes relataram que o contexto de química presente foi envolvente, porém ainda tiveram dificuldades com a dependência do conhecimento prévio. No DOT V3 a consistência interna melhorou e mostrou diferenças significativas de pontuação entre os níveis educacionais, indicando boa validade de critério.

Em conclusão, o desenvolvimento do teste DOT e a sua consequente avaliação ao longo de suas três iterações aprimorou suas capacidades de medição do pensamento crítico em química, e configurou-o como uma ferramenta mais clara na avaliação do pensamento crítico.

Voltando ao Quadro 24, com um olhar para os artigos que avaliaram tecnologias, observamos a repetição do termo *ambiente* como ambiente online, ambiente de aprendizagem ou ambiente digital em 4 artigos: A068, A126, A201 e A051. Esses artigos avaliaram o uso de ambientes digitais. A fim de exemplificar a natureza dos artigos da C04, no que diz respeito aos que avaliaram tecnologias, e em

especial a avaliação de ambientes, devido a sua maior incidência dentro da categoria, apresentamos um excerto do resumo do A068:

Este artigo relata estudos que exploram a eficácia de um ambiente virtual baseado em um laboratório de química como uma ferramenta para preparar estudantes universitários de química que estudam à distância para suas escolas residenciais no campus [...]. Em um estudo experimental, descobriu-se que o ambiente foi capaz de ser eficaz como uma ferramenta para familiarizar os alunos com o laboratório. No entanto, quando o recurso foi fornecido a alunos à distância, menos da metade dos alunos optou por usá-lo, possivelmente devido ao fato de que o uso do recurso não era necessário para a avaliação na disciplina. Os dados do questionário e da entrevista sugeriram que a maioria dos que usaram o recurso considerou que era uma ferramenta preparatória valiosa e recomendaria seu uso posterior. Para muitos alunos, no entanto, a falta de familiaridade com o laboratório não foi vista como a principal fonte de sua ansiedade e, portanto, um recurso que lhes permitiu se familiarizar com o laboratório não teve um grande impacto em sua experiência de aprendizagem (A068, p. 853, tradução nossa)⁷³.

A partir do exposto o A068 artigo relata suas experiências ao fornecer aos alunos um CD-ROM contendo o laboratório virtual de Química em um ambiente 3D simulando uma representação precisa dos laboratórios de ensino. O estudo relatado em A068 avaliou a eficácia do laboratório virtual usando uma combinação de questionários e entrevistas.

Um dos objetivos principais do A068 foi avaliar a eficácia do laboratório virtual para aprender sobre a estrutura espacial do laboratório e os itens de aparelhos e equipamentos contidos nele. Para isso, um design experimental foi usado, com os participantes divididos em dois grupos: um grupo explorou o laboratório virtual; e um grupo foi levado em um tour pelo laboratório real. Os participantes de cada grupo foram então testados em sua lembrança do layout do laboratório e sua familiaridade com os aparelhos. Os autores do A068 definiram que se o aprendizado

⁷³ No texto original: "This article reports on studies exploring the effectiveness of a virtual environment based on a chemistry laboratory as a tool to prepare university chemistry students studying at a distance for their on-campus residential schools, in response to evidence suggesting that many of these students experienced a lack of confidence and a sense of anxiety approaching these sessions. In an experimental study it was found that the environment was able to be effective as a tool for familiarising students with the laboratory. However, when the resource was provided to distance students, less than half of the students chose to use it, possibly due to the fact that use of the resource was not required for the assessment in the subject. Questionnaire and interview data suggested that most of those who used the resource found that it was a valuable preparatory tool and would recommend its further use. For many students, however, a lack of familiarity with the laboratory was not seen as the major source of their anxiety and therefore a resource allowing them to become familiar with the laboratory did not have a major impact on their learning experience" (A068, p. 853).

usando o laboratório virtual foi considerado equivalente ou quase equivalente ao aprendizado dos alunos que realizaram um tour pelo laboratório real, poderia ser concluído que o laboratório virtual é uma ferramenta preparatória eficaz para ganhar familiaridade com o ambiente do laboratório.

De acordo com o A068, os participantes do laboratório real pontuaram em média mais alto do que os participantes do laboratório virtual, porém a diferença não foi significativa no nível de 95% em nenhum dos testes. Logo os resultados sugerem que aprender explorando o laboratório real provavelmente é mais eficaz do que aprender por meio do laboratório virtual, mas a diferença é considerada pequena. Assim, “o laboratório virtual é uma ferramenta eficaz para ganhar familiaridade com o ambiente do laboratório, especialmente para um aluno estudando à distância que não tem a oportunidade de explorar o laboratório antes da escola residencial” (A068, p. 858, tradução nossa)⁷⁴.

Os resultados do A068 também posicionam o laboratório virtual como um recurso para alunos a distância abordarem os problemas identificados de falta de confiança e sensação de ansiedade encontrados suas sessões de laboratório em escolas residenciais. A maioria dos alunos sentiu que o laboratório virtual os ajudou a identificar e localizar itens de aparelhos dentro do laboratório. Nesse sentido 41% indicaram que achavam que isso melhorou sua confiança e 32% indicaram que isso contribuiu para reduzir sua ansiedade.

Os resultados do questionário aplicado aos participantes mostraram que a falta de familiaridade com o laboratório não foi um problema tão grande quanto esperado. Apenas 19% dos alunos acharam difícil identificar itens de aparelhos e 21% acharam difícil localizá-los. No entanto, surgiram problemas com cálculos matemáticos e a aplicação de conceitos de química nas tarefas práticas.

Dado que a capacidade de aplicar técnicas matemáticas e conceitos de química dentro das sessões práticas surgiu como uma grande fonte de ansiedade dos alunos, o A068 sugere-se que a incorporação de instruções para esses aspectos da tarefa pode ser fornecida, e seria um aprimoramento valioso para o ambiente virtual.

⁷⁴ No texto original: “[...] the Virtual Laboratory is an effective tool for gaining familiarity with the laboratory environment, especially for a student studying at a distance who does not have the opportunity to explore the laboratory before the residential school” (A068, p. 858).

A partir do exposto, conclui-se que os ambientes virtuais 3D interativos têm grande potencial educacional devido à sua capacidade de envolver os alunos na exploração, construção e manipulação de objetos virtuais, estruturas e representações. O uso de laboratórios virtuais ou simulações de laboratório possuem potencialidades específicas, como: fornecer aos alunos representações visuais de conceitos de química, preparar os alunos para suas sessões de laboratório e reduzir sensações de ansiedade e falta de confiança quanto às práticas em laboratório real, servindo como uma etapa útil para mitigar esses obstáculos antes da visita ao laboratório real.

4.2.5 Avalia compreensões, percepções e interesse (C05)

A categoria C05 consistiu dos artigos que avaliaram compreensões, percepções e interesse. No Quadro 25 apresentamos todas as subcategorias encontradas na categoria C05 e os respectivos códigos desses artigos. As subcategorias mostram as compreensões, percepções e interesse a respeito dos temas que foram avaliados.

Quadro 25 - C05: Avalia compreensões, percepções e interesse

Subcategoria	Código
Avalia declarações explicativas dos alunos sobre fenômenos macroscópicos de titulações e cromatografia	A017
Avalia explicações dos alunos sobre reações de substituição por meio do uso de técnicas de análise léxica e regressão logística	A018
Avalia a compreensão dos alunos ao longo de uma intervenção em sala de aula de aprendizagem baseada em jogos	A029
Avalia os desenvolvimentos na compreensão do aluno a respeito da linguagem química	A032
Avalia a compreensão autoavaliada dos alunos ao se envolver em atividades instrucionais	A034
Avalia as compreensões dos professores sobre o desenvolvimento da avaliação de habilidades de competência representacional.	A038
Avalia a compreensão do aluno sobre o tema PNM	A065
Avalia as compreensões dos professores de química em formação sobre equilíbrios em soluções ácido-base	A136
Avalia as concepções que os alunos têm sobre diagramas de energia	A176
Avalia as compreensões dos alunos sobre palavras e representações científicas	A177
Avalia as compreensões dos alunos sobre as ideias de química do ensino médio	A186

Avalia as concepções alternativas de professores sobre a bateria de chumbo-ácido, por meio de uma avaliação diagnóstica	A185
Avalia as compreensões e percepções de professores de química do ensino médios a respeito da Alfabetização da Informação e as dificuldades de sua implementação	A091
Avalia as expectativas cognitivas de professores de química sobre o aprendizado de química, ou seja, suas percepções a respeito do que um aluno deve saber em química, e sua influência na construção de questões de exame em química geral	A112
Avalia as compreensões e atitudes dos alunos em química verde ⁷⁵	A055
Avalia a percepção dos alunos sobre a matemática	A013
Avalia a percepção dos alunos sobre as estruturas de Kekulé	A011
Avalia as percepções dos alunos ao fazer um produto digital de <i>blended media</i>	A132
Avalia o interesse de professores de ciências em formação e o tema ácidos e bases	A079

Fonte: os autores

Analisando o Quadro 25, que diz respeito aos artigos que avaliaram as compreensões, percepções e interesse, a sua maioria (68% dos artigos da categoria C05 – 13 artigos) conduziu suas avaliações com alunos. Já os que avaliaram as compreensões, percepções e interesse de professores consistiram em apenas 32% da categoria C05 (6 artigos). Portanto, há uma tendência em priorizar a avaliação de alunos, ao invés de professores, quando se diz respeito às suas compreensões, percepções e interesse.

Para exemplificar os artigos que avaliaram compreensões, percepções e interesse de alunos, trazemos um excerto do artigo A011 abaixo:

Estudantes de química frequentemente demonstram dificuldade com o princípio da ressonância. Apesar de muitas tentativas de mitigar essa dificuldade, houve poucas tentativas de examinar a causa raiz desses problemas. Neste estudo, os alunos foram avaliados quanto à percepção das estruturas de Kekulé com base na teoria da aprendizagem perceptual, que é fundamentada em mecanismos cognitivos de percepção visual. Os dados dessa avaliação mostram que os alunos estão percebendo pistas inapropriadas dessa representação, o que infere que a própria imagem pode ser um impedimento para o aprendizado sobre ressonância (A011, p. 659, tradução nossa)⁷⁶.

⁷⁵ Apesar de também avaliar as atitudes, o artigo apresentou maior ênfase em avaliar as compreensões, verificando que os alunos aprenderam conceitos de química verde e perceberam que a química tem conexões com seus futuros cursos e profissões.

⁷⁶ No texto original: "Students in chemistry often demonstrate difficulty with the principle of resonance. Despite many attempts to mitigate this difficulty, there have been few attempts to examine the root cause of these issues. In this study, students were assessed for their perception of Kekule' structures based on perceptual learning theory, which is grounded in cognitive mechanisms of visual perception. The

Nota-se que o artigo A011 avalia as percepções dos alunos em relação às estruturas de Kekulé⁷⁷. No artigo, as estruturas de Kekulé foram empregadas para compreender o conceito de ressonância. Comumente, os alunos possuem dificuldade em compreender a ressonância, em parte devido a interpretações errôneas e desafios na transição entre diferentes representações químicas. Isso ocorre, pois a competência representacional é frequentemente subdesenvolvida nos alunos, levando a tais dificuldades. A competência representacional envolve entender representações macroscópicas, submicroscópicas e simbólicas e traduzir pontos ou elementos-chave entre elas.

A pergunta de pesquisa que moveu o artigo A011 foi: “Quais características/acessibilidades os alunos percebem ao visualizar uma representação canônica de ressonância e como essa percepção afeta sua conceituação de ressonância?”. Para responder a essa pergunta, foi necessário avaliar as percepções existentes dos alunos a respeito da ressonância.

O estudo envolveu 33 alunos matriculados em uma disciplina de Química Geral para alunos dos cursos de Bioquímica e Química. A disciplina usou a metodologia *Process Oriented Guided Inquiry Learning* (POGIL), ou *Aprendizagem de investigação guiada e orientada em processos*, em português. Os alunos se envolveram com estruturas de ressonância do benzeno e, na sequência, responderam a uma questão de Reflexão sobre Aprendizagem a respeito de representações de ressonância. Suas respostas foram analisadas e as suas percepções avaliadas. Após as análises, conceitos comuns foram identificados entre as percepções dos alunos, como, por exemplo, a rotação de ligação, simetria e elétrons deslocalizados. Esses conceitos foram agrupados em códigos e posteriormente categorizados.

A respeito dos resultados do A011, a maioria dos alunos (78%) viu as estruturas hexagonais como entidades distintas em vez de representações de um único fenômeno. Isso demonstrou um mal-entendimento persistente, apesar da

data from this assessment shows that students are perceiving inappropriate clues from this representation, which infers that the image itself might be an impediment to learning about resonance” (A011, p. 659).

⁷⁷ Estruturas de Kekulé são uma maneira de representar ligações em compostos orgânicos, particularmente em moléculas que contêm benzeno e outros compostos aromáticos. Essas representações ajudam na visualização da natureza deslocalizada dos elétrons em sistemas aromáticos, onde os elétrons não são fixados entre ligações específicas, mas são compartilhados entre vários átomos (Atkins; Friedman, 2011).

instrução POGIL realizada anteriormente na disciplina. A análise das respostas dos alunos revelou que os alunos frequentemente usavam Dimensionalização e Segmentação, focando em diferenças visuais em vez do conceito unificado de ressonância. A unitização, que se alinha com a ideia de ressonância, foi observada com menos frequência, indicando uma lacuna significativa na compreensão dos alunos.

Assim, os autores do A011 concluem que os alunos frequentemente veem as estruturas de Kekulé como duas espécies distintas em vez de representações de uma única molécula. Isso cria modelos mentais persistentes e resistentes à mudança. A consequência disso é que os alunos atribuem, inadequadamente, distinção a características que deveriam ser vistas como juntas. Para abordar esses equívocos, o A011 exemplifica a aplicação de uma abordagem metarrepresentacional, em um segundo momento do artigo. Conforme as discussões dos autores, essa abordagem permitiu que os alunos analisassem a representação de modo mais crítico.

Semelhante ao artigo A011, os outros artigos alocados a categoria C05 também avaliaram as percepções, compreensões e interesse de alunos e professores a respeito de diferentes temas e conceitos que envolvem aspectos da Química. Nesse sentido, na sequência fazemos uma divisão dos artigos que avaliaram compreensões, percepções e interesse de professores e alunos.

Em relação aos artigos que avaliaram percepções, compreensões e interesse de professores, esses foram a respeito de: avaliação de habilidades de competência representacional (A038); equilíbrios em soluções ácido-base (A136); bateria de chumbo-ácido (A185); Alfabetização da Informação (A091); aprendizado de química (A112); e ácidos e bases (A079).

Em relação aos artigos que avaliaram percepções, compreensões e interesse de alunos, esses foram a respeito de: fenômenos macroscópicos de titulações e cromatografia (A017); reações de substituição (A018); compreensões ao se envolverem em uma intervenção de aprendizagem baseada em jogos (A029); linguagem química (A032); compreensões ao se envolverem em atividades instrucionais (A034); o tema PNM (A065); diagramas de energia (A176); palavras e representações científicas (A177); ideias de química do ensino médio (A186); química verde (A055); matemática (A013); estruturas de Kekulé (A011); e percepções ao se envolverem na criação de um produto digital de blended media (A132).

A fim de continuarmos a explicitação dos artigos alocados na categoria C05, trazemos um excerto do A079 abaixo:

Os alunos têm mais probabilidade de ter sucesso em tópicos de seu interesse do que em outros. Este estudo tem como objetivo desenvolver uma Escala de Interesse Ácido-Base (ABIS) e, posteriormente, avaliar o interesse de futuros professores de ciências em ácidos-bases de acordo com o gênero, anos na universidade, tipo de ensino médio que os futuros professores de ciências frequentaram e seu sucesso relativo no módulo Química Geral II. Após cuidadosa consideração da validade e confiabilidade para o desenvolvimento da ABIS, os dados foram coletados de 453 futuros professores de ciências. A ABIS do tipo Likert de cinco pontos consistiu em 26 itens e envolveu três fatores (“Interesse Individual”, “Interesse em Informações Teóricas” e “Interesse relacionado à Vida Diária”). O coeficiente alfa de Cronbach que representa a confiabilidade da ABIS foi de 0,894. Uma vez que a confiabilidade da ABIS foi verificada, ela foi administrada a 982 futuros professores de ciências em oito universidades públicas na Turquia (A079, p. 630, tradução nossa)⁷⁸.

A partir do exposto, observa-se que o A079 desenvolveu uma escala de interesse ácido-base e na sequência avaliou o interesse de licenciandos a respeito de ácidos-bases. Conforme o A079, o interesse é compreendido como a conexão de um indivíduo com objetos ou situações, que consequentemente influencia seu engajamento e satisfação em atividades (Krapp, 2002; Krapp, 2005). O A079 argumenta que pesquisas que investigam interesse no contexto educacional são relevantes, pois o interesse impulsiona a participação dos alunos em atividades realizadas dentro e fora da sala de aula, levando a um melhor desempenho em assuntos que os alunos consideram interessantes.

Para o A079, o interesse pode ser categorizado em interesse situacional, uma resposta de curto prazo a estímulos ambientais, e interesse individual, uma inclinação pessoal de longo prazo moldada por características e experiências (Krapp, 2002). O interesse individual se desenvolve por meio da

⁷⁸ No texto original: “Students are more likely to be successful in topics they are interested in than others. This study aims to develop an Acid–Base Interest Scale (ABIS) and subsequently evaluate the interest of pre-service science teachers in acids–bases according to gender, years at the university, type of high school the pre-service science teachers attended, and their relative success in the module General Chemistry II. Upon careful consideration of validity and reliability for the development of the ABIS, data were collected from 453 pre-service science teachers. The five-point Likert-type ABIS consisted of 26 items and entailed three factors (“Individual Interest”, “Interest in Theoretical Information”, and “Interest related to Daily Life”). Cronbach’s alpha coefficient representing the reliability of the ABIS was 0.894. Once the reliability of the ABIS was ascertained, it was administered to 982 pre-service science teachers in eight public universities in Turkey” (A079, p. 630).

repetição e promove a curiosidade, a dedicação e a atenção em contextos educacionais (Hidi; Renninger, 2006).

A respeito do método, o estudo foi realizado em duas etapas: o desenvolvimento de uma Escala de Interesse Ácido-Base (ABIS) e a avaliação do interesse de licenciandos em ciências sobre ácidos e bases. Para as discussões apresentadas na tese, concentramo-nos nas discussões a respeito da segunda etapa, pois esta justificou a alocação do A079 para a categoria C05.

Para a avaliação do interesse dos licenciandos a respeito de ácidos e bases, 982 licenciandos de várias universidades públicas na Turquia foram investigados. Os dados foram analisados usando software estatístico, com foco na confiabilidade (alfa de Cronbach) e validade (análise fatorial confirmatória). Além disso, um teste t de amostra independente avaliou as diferenças de gênero, enquanto a ANOVA unidirecional avaliou o interesse com base no ano universitário, tipo de ensino médio e notas na disciplina de Química Geral II.

A respeito dos resultados, O ABIS apresentou uma pontuação média de interesse de 94,313, indicando medições confiáveis. O A079 encontrou diferenças significativas de interesse entre gêneros, já que as licenciandas mostraram maior interesse do que os licenciandos. Também percebeu uma variação do interesse conforme o ano de curso, atingindo um pico no segundo ano. A respeito do tipo de ensino médio dos licenciandos e suas notas na disciplina de Química Geral II, não foi encontrado diferença significativa. Finalmente observou-se um decréscimo em interesse ao longo da progressão acadêmica. O A079 traça discussões que relacionam esse achado com o aumento da exposição aos assuntos ao decorrer do curso.

Os resultados do A079 são relevantes, pois destacam a necessidade de estimular e sustentar o interesse em conceitos centrais da disciplina de Química ao longo do curso de Licenciatura. Essa necessidade merece destaque, pois o interesse dos estudantes pode influenciar diretamente sua motivação e desempenho acadêmico. A formação inicial de professores apresenta desafios, dado que muitos alunos podem perceber certas disciplinas como abstratas ou difíceis de relacionar com suas vidas cotidianas. Logo, é importante que os formadores de professores implementem estratégias didáticas que tornem os conceitos químicos mais acessíveis, relevantes e interessantes.

4.2.6 Avalia conhecimentos (C06)

A categoria C06 consistiu dos artigos que avaliaram conhecimentos. No Quadro 26 apresentamos todas as subcategorias encontradas na categoria C06 e os respectivos códigos desses artigos. As subcategorias mostram quais conhecimentos foram avaliados.

Quadro 26 - C06: Avalia conhecimentos

Subcategoria	Código
Avalia os desafios dos alunos e processos de solução para o conceito de mol e o conceito de estequiometria	A012
Avalia o conhecimento conceitual e algorítmico dos alunos em química geral com exames ACS	A023
Avalia o desenvolvimento do conhecimento de professores quando expostos à aprendizagem baseada em contexto	A024
Avalia e revela os algoritmos e procedimentos memorizados nos quais os alunos confiaram para responder às perguntas	A043
Avalia o conhecimento dos estudantes. Os resultados do artigo mostraram que os participantes foram fracos em responder a problemas de fator de conversão, computacionais e estequiométricos	A053
Avalia e compara a qualidade das estruturas cognitivas de alunos do ensino médio e professores de química em formação no domínio da química orgânica.	A064
Avalia o conhecimento pedagógico do conteúdo de professores	A084
Avalia o conhecimento pedagógico do conteúdo de professores	A085
Avalia o desempenho acadêmico em química baseado no conhecimento de conceitos, aplicação e a resolução de problemas	A066
Avalia como os alunos fizeram conexões entre reações em um novo currículo com uma tarefa de ordenação de cartas.	A129
Avalia o conhecimento pedagógico do conteúdo de licenciandos a respeito da Natureza da Ciência	A135
Avalia como os alunos utilizam o conhecimento do conteúdo químico e o aplicam ao resolver problemas de química baseados em contexto	A154
Avalia a natureza específica do tópico do conhecimento pedagógico do conteúdo de dois professores de química experientes	A173
Avalia as estruturas cognitivas dos alunos sobre os conceitos de solução e dissolução	A181
Avalia as estruturas de conhecimento a respeito de conceitos de ligação de eletronegatividade e ligações covalentes polares	A183
Avalia as estruturas de conhecimento dos alunos	A188
Avalia os alunos no conceito de posição de partícula durante a mudança física reversível de dissolução	A182
Avalia a aprendizagem dos alunos com uma unidade de Química Conectada, um ambiente baseado em computador	A147
Avalia o conhecimento dos futuros professores sobre compostos aromáticos	A010

Avalia os acertos, estratégias e erros comuns dos alunos em suas respostas a perguntas que envolviam o formalismo de empurrar elétrons (seta curva) (EPF), parte da linguagem da química orgânica.	A114
--	------

Fonte: os autores

Ao analisarmos o Quadro A025, notamos uma prevalência em avaliar os conhecimentos de alunos ao invés de avaliar os conhecimentos de professores, uma vez que 70% dos artigos da categoria C06 avaliaram os conhecimentos dos alunos.

Em relação aos conhecimentos avaliados na categoria C06 na totalidade, estes foram a respeito de: o conceito de mol e o conceito de estequiometria; o conhecimento conceitual e algorítmico em química geral; os conhecimentos de professores quando expostos à aprendizagem baseada em contexto; conhecimentos algoritmos e procedimentos memorizados; conhecimentos de conversão, computacionais e estequiométricos; conhecimentos da química orgânica; o conhecimento pedagógico do conteúdo; o conhecimento de conceitos, aplicação e a resolução de problemas; o conhecimento sobre conexões entre reações; o conhecimento do conteúdo químico e sua aplicação ao resolver problemas de química baseados em contexto; solução e dissolução; conceitos de ligação de eletronegatividade e ligações covalentes polares; estruturas de conhecimento; a dissolução; Química Conectada; compostos aromáticos; e a linguagem da química orgânica.

Já os artigos que avaliaram os conhecimentos de professores e futuros professores, estes compuseram 30% da categoria C06. Dentre esse grupo de artigos, observou-se a tendência em avaliar o conhecimento pedagógico do conteúdo de professores, já que essa característica foi repetida nos artigos: A084, A085, A135 e A173.

Trazemos um excerto do artigo A085 para exemplificar os artigos alocados para a categoria C06:

Este artigo teve como objetivo explorar como o conhecimento pedagógico do conteúdo (PCK)⁷⁹ dos professores influenciou suas adaptações dos materiais curriculares do novo currículo de química do ensino médio, um currículo de ciências baseado em padrões, na China. [...] Uma abordagem interpretativa foi adotada e cinco professores de química experientes em quatro escolas de ensino

⁷⁹ Pedagogical content knowledge.

médio participaram deste estudo. Observações em sala de aula e entrevistas foram usadas como métodos de pesquisa.

Logo o artigo A085 avaliou o conhecimento pedagógico do conteúdo de cinco professores de química experientes (mais de dez anos de experiência no ensino) dos Anos Finais do Ensino Fundamental⁸⁰ na China. O enfoque do artigo esteve em compreender como o conhecimento pedagógico do conteúdo influenciou suas adaptações dos materiais curriculares do currículo. Para análise do conhecimento pedagógico do conteúdo dos professores, utilizou-se o modelo de Conhecimento Pedagógico de Conteúdo de Magnusson *et al.* (1999).

A respeito da coleta de dados, a pesquisa foi focada em uma unidade curricular inteira para obter uma imagem completa dos materiais curriculares e da prática do professor. Os materiais curriculares observados, foram divididos em unidades, seções e módulos. Para cada unidade, os objetivos de ensino, estratégias de ensino foram selecionados como os componentes do currículo pretendido. Para identificar as adaptações dos professores do currículo pretendido, e os vários componentes do conhecimento pedagógico que levaram a essas adaptações, observações em sala de aula e entrevistas semiestruturadas foram realizadas. As observações foram realizadas durante toda unidade (duas a quatro semanas). Para a análise, utilizou-se a abordagem interpretativa.

Os resultados do artigo A085 revelam que haviam quatro componentes do conhecimento pedagógico do conteúdo que levaram às adaptações dos materiais curriculares pelos professores. Estes foram: o conhecimento da compreensão dos alunos sobre ciências, o conhecimento da avaliação em ciências, o conhecimento de estratégias de ensino e o conhecimento do currículo de ciências. Entre esses quatro componentes, o conhecimento da compreensão dos alunos sobre ciências e o conhecimento da avaliação em ciências foram os fatores que mais influenciaram as adaptações dos materiais curriculares pelos professores. Apesar disso, os resultados são específicos para o contexto investigado, professores experientes na China, e podem não ser generalizáveis para outras regiões, especialmente devido às particularidades na formação de professores da região e o currículo. As discussões do artigo A085 revelam a importância de incorporar o conhecimento pedagógico do conteúdo de professores experientes no design do

⁸⁰ No texto original, professores de “Middle School”.

currículo e sugere o envolvimento de professores experientes e menos experientes para melhorar a eficácia do currículo.

4.2.7 Avalia o currículo (C07)

A categoria C07 consistiu dos artigos que avaliaram o currículo. De acordo com Lopes e Macedo (2014) responder à pergunta simples “o que é currículo” não é uma tarefa fácil. Isso é devido às diferentes definições que o currículo recebeu ao longo da história e aos diferentes contextos escolares em que o currículo é discutido. Contudo há um aspecto comum: a ideia de organização, prévia ou não, de experiências e/ou situações de aprendizagem realizadas por docentes e/ou redes de ensino com o intuito de levar a cabo um processo educativo (LOPES; MACEDO, 2014).

Assim, baseamo-nos na definição de Lopes e Macedo (2014), as sete acepções dos dicionários: *Houaiss* e o *The Greenwood Dictionary of Education* apresentados anteriormente no Quadro 20. Evidentemente, também houveram artigos que utilizaram o próprio termo “currículo” em seus objetivos ou questões, facilitando a categorização.

No Quadro 27 apresentamos todas as subcategorias encontradas na categoria C07: Avalia o currículo e os respectivos códigos desses artigos. As subcategorias mostram quais currículos foram avaliados.

Quadro 27 - C07: Avalia o currículo

Subcategoria	Código
Avalia e compara documentos para o bacharel em química dos EUA, Europa e Austrália.	A044
Avalia como os Learning Outcomes ⁸¹ [objetivos educacionais] são ensinados, praticados e avaliados em química orgânica	A076
Avalia os esforços de reforma curricular implementados em um curso fundamental de pós-graduação	A095
Avalia a disciplina de físico-química no contexto nacional	A125
Avalia o impacto de um currículo orientado para a discussão	A014
Avalia um programa de formação de professores e como ele afetou a realização das avaliações formativos dos professores de formação inicial	A137
Avalia o treinamento que assistentes de professores recebem por meio da análise de suas percepções a respeito do treinamento	A157
Avalia conteúdos curriculares de química na República Tcheca, Finlândia e Turquia foram analisados, e seus objetivos foram comparados	A168

⁸¹ Programação parcial de um curso ou matéria.

Avalia a “combinação” entre o status piagetiano dos alunos do ensino médio de Botswana e as demandas cognitivas da educação em química. curso de química (ementas e sequenciamento de tópicos)	A171
Avalia o novo design de um curso baseado puramente na conclusão de uma série de desafios.	A174
Avalia os currículos de química da Universidade de Cardiff	A200
Avalia materiais didáticos para melhorar a compreensão dos ciclos biogeoquímicos	A031
Avalia um projeto de aprendizagem de serviço interativo	A036

Fonte: os autores

Como pode ser observado no Quadro 27, os artigos da categoria C07 avaliaram uma variedade de currículos em diferentes contextos. Trazemos abaixo um excerto do artigo A044 para exemplificar a natureza dos artigos alocados na categoria C07:

Em um projeto da International Union of Pure and Applied Chemistry (IUPAC), desenvolvemos um método de uso da avaliação crítica de resultados de aprendizagem e descritores em níveis internacional, nacional e institucional para aprimorar a educação em química centrada no aluno. Este método é focado no processo, visando aprender por meio do compartilhamento e comparação de práticas ao redor do mundo. Três documentos abrangentes para o bacharelado em química dos EUA, Europa e Austrália foram comparados. As diferenças foram encontradas mais no estilo do que no conteúdo. Uma ferramenta de autoanálise foi construída para avaliar como os resultados de aprendizagem para cursos e módulos estão vinculados entre si e aos resultados de aprendizagem para programas educacionais e como os resultados de aprendizagem esperados podem ser alinhados com atividades de aprendizagem e avaliação (A044, p. 427, tradução nossa)⁸².

A partir do exposto, identificamos que o artigo A044 avalia e compara documentos para o bacharel em química dos EUA, Europa e Austrália. De acordo com os autores, o ensino de Química focado em Resultados de Aprendizagem (*learning outcomes*, no inglês) está sendo cada vez mais praticado, fornecendo novas oportunidades para comparações internacionais. Esse aumento do interesse nos Resultados de Aprendizagem se dá devido às pesquisas que defendem o seu uso no

⁸² No texto original: “In an International Union of Pure and Applied Chemistry (IUPAC) project, we have developed a method of using critical evaluation of learning outcomes and descriptors at international, national, and institutional levels to enhance learner-centered chemistry education. This method is process-focused, aimed at learning by sharing and comparing practices around the world. Three overarching documents for the chemistry bachelor from the USA, Europe and Australia were compared. The differences were found to be more in style than in content. A tool for self-analysis was constructed to evaluate how learning outcomes for courses and modules are linked to each other and to learning outcomes for educational programs and how the expected learning outcomes can be aligned with learning activities and assessment” (A044, p. 427).

ensino superior (Biggs; Tang, 2011) e decisões políticas. Um exemplo desse último é o Processo de Bologna (Pinto, 2010), uma série de reuniões ministeriais e acordos entre as instituições europeias que ocorreram entre 2001-2020, que culminou na criação do *European Higher Education Area* (EHEA), projetado para garantir sistemas de ensino superior mais comparáveis, compatíveis e coerentes na Europa.

A respeito dos Resultados de Aprendizagem, o A044 os define como:

[...] declarações do que se espera que um aluno saiba, entenda e/ou seja capaz de demonstrar após a conclusão de um processo de aprendizagem. Um exemplo hipotético de um resultado de aprendizagem pretendido poderia ser: o aluno deve ser capaz de interpretar dados obtidos de observações e medições de laboratório em termos de sua significância e relacioná-los à teoria apropriada (A044, p. 427, tradução nossa)⁸³.

Assim, defende-se que o Ensino de Química apoiado em Resultados de Aprendizagem, levaria a uma educação em química mais centrada no aluno, com os alunos se tornando aprendizes mais ativos. Os Resultados de Aprendizagem são delineados em vários níveis, tanto em documentos internacionais como em documentos nacionais, bem como em documentos mais locais, como em currículos institucionais e programas de cursos individuais

A pesquisa do artigo A044 compara os Resultados de Aprendizagem em duas dimensões: entre documentos em níveis semelhantes de diferentes regiões e em níveis diferentes em um sistema. Vale destacar que o objetivo não é prescrever uma estrutura específica com esse movimento, mas sim enfatizar a importância da consistência nos Resultados de Aprendizagem em diferentes níveis para promover uma aprendizagem eficaz do aluno. O artigo A044 também examina o alinhamento dos Resultados de Aprendizagem com atividades e avaliações no nível do curso/módulo.

Um método focado no processo foi usado para avaliar e fazer as comparações internacionais. Esse processo envolveu dois workshops com educadores e pesquisadores de química de vários países. O primeiro workshop se concentrou em projetar uma matriz para comparar documentos abrangentes dos EUA,

⁸³ No texto original: “[...] statements of what a learner is expected to know, understand and/or be able to demonstrate after completion of a process of learning. A hypothetical example of an intended learning outcome could be: The student should be able to interpret data obtained from laboratory observations and measurements in terms of their significance and relate them to appropriate theory” (A044, p. 427).

Europa e Austrália. Essa matriz, que avaliou habilidades e competências essenciais, foi usada como uma ferramenta de autoanálise no segundo workshop.

Os educadores e pesquisadores do workshop analisaram documentos de diferentes níveis e discutiram seu conteúdo, estilo e ênfase. Os workshops promoveram discussões e análise críticas entre os participantes, e ajudaram a refinar o processo de comparação de Resultados de Aprendizagem. Após esse movimento, a análise continuou nos níveis de graduação e curso/módulo, usando a ferramenta de autoanálise para avaliar itens como alinhamento e consistência.

A respeito dos resultados do artigo A044, os documentos abrangentes sobre programas de bacharelado em Química da Europa, EUA e Austrália, inicialmente parecem bem distintos entre si, porém um olhar mais atento revela que as diferenças são mais estilísticas do que substantivas.

O *Eurobachelor* e os *Australian Learning and Teaching Academic Standards* (LTAS), Padrões Acadêmicos de Ensino e Aprendizagem Australianos em português, apresentam seus Resultados de Aprendizagem separadamente. Já as diretrizes do ACS dos Estados Unidos misturam seus Resultados de Aprendizagem com outros conteúdos, como desafios futuros ou atividades de aprendizagem. Por exemplo, as diretrizes do ACS enfatizam o trabalho em equipe e a liderança, enquanto os padrões australianos declaram claramente os Resultados de Aprendizagem como “Declarações de Resultados de Aprendizagem Limiares”. O *Eurobachelor* descreve as competências como aspirações.

Para facilitar a comparação, utilizou-se um esquema unificado em que os Resultados de Aprendizagem foram categorizados em: Habilidades Científicas; Habilidades de Laboratório; Habilidades de Literatura e Literacia Química; Habilidades de Comunicação e Interação; e Responsabilidade Pessoal e Social. De acordo com o A044, apesar de haver diferenças entre os documentos, as competências essenciais contidas nelas são consideravelmente semelhantes. A respeito das distinções, trazemos as mais notáveis na sequência.

Habilidades Científicas: os documentos dos EUA e da Austrália enfatizam a definição de problemas e a formulação de hipóteses, enquanto o *Eurobachelor* se concentra na resolução de problemas e análise. Habilidades de Laboratório: o *Eurobachelor* fornece descrições detalhadas, incluindo instrumentação e documentação, enquanto as diretrizes do ACS se concentram na segurança do

laboratório e no tratamento de emergências. Manuseio de informações: as diretrizes do ACS e do *Eurobachelor* enfatizam a importância da literatura científica, enquanto o LTAS enfatiza a síntese de informações de várias fontes. Comunicação: o LTAS exige diversas habilidades de comunicação e habilidades em línguas estrangeiras, enquanto as diretrizes do ACS incluem liderança no trabalho em equipe. Responsabilidade: todos os documentos enfatizam a importância da responsabilidade. O *Eurobachelor* menciona brevemente o compromisso ético, enquanto o LTAS inclui expectativas abrangentes de responsabilidade. As diretrizes do ACS e o LTAS também abordam o papel global da química e do aprendizado autodirigido.

A respeito das conclusões do artigo A044, verificou-se que o processo de análise e comparação de sistemas educacionais foi tão valioso quanto a lista final de similaridades e diferenças. Isso porque, o processo ampliou as perspectivas dos participantes sobre os currículos dos cursos e programas. Assim, recomenda-se um “método focado em processo para a comparação de documentos educacionais internacionais” (A044, p. 432). Isso é válido, pois o método mostrou potencial para comparar sistemas educacionais e alinhar os resultados da aprendizagem com atividades e avaliações, o que pode levar ao aprimoramento dos programas educacionais. O método incentiva o diálogo entre educadores e pesquisadores sobre os objetivos da educação em química e apoia uma abordagem comparativa que respeita as diversas necessidades educacionais sem impor um único “padrão ouro”. Isso, pois o A044 também destaca a importância da diversidade nos currículos para atender as particularidades do contexto local.

Por fim, embora esses documentos ofereçam insights valiosos, eles refletem principalmente perspectivas da Europa, dos EUA e da Austrália. Desse modo, há necessidade de discussões que incluam economias em crescimento e países em desenvolvimento para refinar e adaptar competências para um entendimento global.

Trazemos abaixo um excerto do A014 para continuarmos a exploração dos artigos pertencentes a categoria C07:

Estratégias instrucionais que dão suporte ao aprendizado significativo dos alunos sobre tópicos químicos complexos são um aspecto importante para melhorar a educação em química. A avaliação adequada do sucesso dessas abordagens pode ser apoiada com o uso de instrumentos alinhados com psicométrica estabelecida. Aqui, relatamos a implementação e avaliação de um desses currículos,

Pensamento Químico, sobre as concepções de acidez relativa dos alunos do primeiro ano de química geral usando o inventário de conceitos recentemente desenvolvido, ACIDI. (A014, p. 543, tradução nossa)⁸⁴

Observamos que o A014 avalia a implementação de um currículo, o Pensamento Químico, com alunos do primeiro ano de Química Geral. Para isso, o A014 utilizou o Inventário Ácido-Base (ACIDI) para analisar as conversas dos alunos, identificar suas concepções e observar o seu desenvolvimento. Assim, o A014 busca estabelecer relações entre o design do currículo, as interações dos alunos e suas compreensões conceituais.

O currículo “Pensamento Químico”, desenvolvido por Talanquer e Pollard (2010) organiza o conteúdo em torno de perguntas fundamentais. Essas promovem discussões que contribuem para uma compreensão mais profunda que vai além da mera memorização de conceitos.

O estudo do A014 foi realizado ao longo de dez semanas em uma grande universidade de pesquisa no nordeste dos Estados Unidos em 2016 e envolveu 230 alunos de primeiro ano em um curso de química geral e orgânica. Os temas abordados dentro do currículo foram: estabilidade termodinâmica, cinética química, equilíbrio, reatividade e eletroquímica.

Uma parte significativa do curso focou em ácidos e bases. A estrutura do curso incluía uma grande aula conduzida pelo professor, seguido de um workshop de aprendizagem cooperativa. As aulas promoviam discussões, com perguntas feitas pelo professor que incentivavam os alunos a responder publicamente e a avaliar as ideias dos colegas, criando um ambiente de discordância cognitiva. Já os workshops, permitiam que os alunos discutissem o material em pequenos grupos.

O ACIDI⁸⁵ foi utilizado como ferramenta diagnóstica antes e depois da instrução. Além disso, utilizou-se um teste postergado aplicado dez semanas após o

⁸⁴ No texto original: “Instructional strategies that support meaningful student learning of complex chemical topics are an important aspect of improving chemistry education. Adequately assessing the success of these approaches can be supported with the use of aligned instruments with established psychometrics. Here, we report the implementation and assessment of one such curriculum, Chemical Thinking, on first-year general chemistry students’ conceptions of relative acidity using the recently-developed concept inventory, ACIDI” (A014, p. 543).

⁸⁵ O ACIDI serviu como ferramenta diagnóstica para identificar as concepções dos alunos sobre a força dos ácidos orgânicos. A avaliação continha nove questões de múltipla escolha, algumas em formato multi-nível. A sequência de três perguntas foi repetida para três conjuntos de compostos: os alunos identificavam o composto mais ácido, justificavam sua escolha, classificavam os outros dois compostos e escolhem a resposta que melhor apoiava sua classificação.

curso. Alguns workshops também foram gravados em vídeo e áudio para analisar as estruturas de argumentação dos alunos.

A respeito dos resultados, notou-se que após a instrução, as pontuações ACIDI melhoraram significativamente do pré-teste (média de 2,59) para o pós-teste (média de 3,74). Além disso, as pontuações pós-teste tardias indicaram boa retenção dessas novas concepções dois meses e meio depois. A respeito das entrevistas, essas revelaram concepções variadas de ressonância. Enquanto alguns alunos reconheceram que grupos funcionais sozinhos não determinam a acidez, outros tiveram dificuldade em compreender como a ressonância impacta a estabilidade. Durante as entrevistas muitos alunos citaram os workshops como benéficos para fornecer uma maior exposição a diferentes abordagens de raciocínio. A análise das discussões sugeriu que o diálogo colaborativo realizado pelos alunos pode promover melhor compreensão sobre os conceitos de ácidos e bases.

Em conclusão, o A014 demonstra que um currículo orientado para a discussão dos alunos melhora a sua compreensão sobre química ácido-base. Os resultados e posteriores discussões do A014 defendem a introdução desse design de currículo de maneira mais precoce nos cursos de graduação a fim de não só aumentar a compreensão dos alunos em temas particularmente desafiadores, mas aprimorar sua capacidade de argumentação sobre esses temas.

4.2.8 Avalia atitudes (C08)

A categoria C08 consistiu dos artigos que avaliaram atitudes. No Quadro 28 apresentamos todas as subcategorias encontradas na categoria C08 e os respectivos códigos desses artigos. As subcategorias mostram quais atitudes foram avaliadas exatamente.

Quadro 28 - C08: Avalia atitudes

Subcategoria	Código
Avalia atitudes emocionais e intelectuais dos alunos em relação à química como disciplina	A063
Avalia atitudes positivas em relação à química	A119
Avalia e analisa as “melhores práticas” ⁸⁶ para avaliações do tipo <i>inquiry</i> relatadas pelos próprios professores	A037

⁸⁶ Destacamos que não acreditamos que existem “melhores práticas” únicas e que estas dependem do contexto em questão. Contudo esses artigos avaliam as práticas relatadas/consideradas pelos próprios professores como “melhores”.

Avalia e analisa as práticas de avaliação e os critérios de avaliação de futuros professores na adequação de vários procedimentos científicos	A138
Avalia estilos de liderança ⁸⁷ em organizações estudantis que participam de extensão de química	A009
Avalia como os professores usam a avaliação para orientar o ensino	A096
Avalia os processos de geração de itens de avaliação de professores de química do ensino médio	A111
Avalia as características dos professores em relação a elaboração de avaliações formativas de química e interpretação dos resultados	A035
Avalia as práticas de avaliação de instrutores de química para a avaliação e classificação de diferentes respostas dos alunos a questões conceituais em avaliações somativas de química geral	A190
Avalia as “melhores práticas” consideradas por professores de química para desenvolver avaliações formativas e as avaliações que eles implementam em seus cursos	A193

Fonte: os autores

Ao analisar o Quadro 28, nota-se uma prevalência em avaliar atitudes de professores, uma vez que 70% dos artigos tiveram essa característica. Esses artigos possuem uma segunda semelhança, eles avaliaram atitudes de professores frente a avaliação que eles realizavam em sala de aula, especificamente suas práticas de avaliação. Logo existe a tendência em avaliar práticas de avaliação ou práticas avaliativas de professores.

Trazemos abaixo um excerto do artigo A193 para exemplificar algumas características dos artigos alocados a categoria C08:

Mesmo quando as crenças dos professores de química sobre o design da avaliação se alinham com as melhores práticas citadas na literatura, barreiras podem impedir que os professores promulguem essas crenças ao desenvolver avaliações do dia a dia. Neste artigo, a relação entre as “melhores práticas” autogeradas pelos professores de química do ensino médio para desenvolver avaliações formativas e as avaliações que eles implementam em seus cursos são examinadas. Os resultados de uma avaliação detalhada de várias avaliações formativas de química do ensino médio, metas de aprendizagem e atividades de aprendizagem revelam que os itens de avaliação são frequentemente desenvolvidos para exigir tarefas bem articuladas, mas carecem de alinhamento em relação ao nível representacional ou empregam apenas um nível representacional para quase todos os itens de avaliação. Implicações para o desenvolvimento de um método específico de química para avaliar o alinhamento são apresentadas,

⁸⁷ O artigo foi alocado a categoria “Avalia atitudes” por conta do significado do termo *liderança*. De acordo com o dicionário é a ação de liderar um grupo de pessoas ou uma organização.

bem como implicações para o design de avaliação de química do ensino médio (A193, p. 457, tradução nossa)⁸⁸.

A partir do exposto, nota-se que o A193 avalia as “melhores práticas” consideradas pelos professores de química para desenvolver avaliações formativas e as avaliações que eles realmente implementam em seus cursos. As perguntas de pesquisa que orientam o estudo foram: (1) Como as tarefas e os níveis representacionais de itens de avaliação formativa planejados podem ser diagramados para permitir a comparação com materiais instrucionais correspondentes (por exemplo, objetivos de aprendizagem e atividades de aprendizagem)? e (2) Qual é a natureza da relação entre as “melhores práticas” autogeradas pelos professores para avaliações formativas e as avaliações que eles geram?

Para abordar as perguntas de pesquisa, as avaliações, atividades de aprendizagem e metas de aprendizagem de professores de química do Ensino Médio foram coletadas. Três professores participaram de um programa de desenvolvimento profissional no ano de 2018, descrito no estudo do A193, que serviu como contexto da pesquisa. Esse programa, foi elaborado com o objetivo de melhorar as práticas de avaliação dos professores. Como parte do desenvolvimento profissional do programa, os professores foram solicitados a fornecer uma avaliação formativa planejada para uma aula, bem como as metas de aprendizagem e atividades de aprendizagem associadas à aula.

Para o A193 as metas de aprendizagem são as expectativas dos alunos definidas pelos professores no início da aula. Todos os professores registraram as metas de aprendizagem nos materiais fornecidos aos alunos no início da aula. Já as atividades de aprendizagem, são os materiais instrucionais usados pelos professores ao longo da aula.

A primeira questão de pesquisa explora como diagramar tarefas e níveis representacionais de avaliações formativas para avaliar seu alinhamento com

⁸⁸ No texto original: “Even when chemistry teachers’ beliefs about assessment design align with literature-cited best practices, barriers can prevent teachers from enacting those beliefs when developing day-to-day assessments. In this paper, the relationship between high school chemistry teachers’ self-generated “best practices” for developing formative assessments and the assessments they implement in their courses are examined. Results from a detailed evaluation of several high school chemistry formative assessments, learning goals, and learning activities reveal that assessment items are often developed to require well-articulated tasks but lack either alignment regarding representational level or employ only one representational level for nearly all assessment items. Implications for the development of a chemistry-specific method for evaluating alignment are presented as well as implications for high school chemistry assessment design” (A193, p. 457).

a instrução. O movimento realizado para responder essa pergunta, gerou um gráfico de alinhamento usado para visualizar recursos de avaliação para aulas de química do ensino médio. Esta ferramenta ajuda os professores a avaliar o alinhamento entre suas avaliações e instruções, oferecendo uma alternativa potencialmente mais fácil do que outras ferramentas mais complexas como o EQuIP. O EQuIP é uma rubrica que avalia sequências de aulas e unidades para verificar seu alinhamento com o NGSS (Next Generation Science Standards). A rubrica usa três domínios: um design tridimensional, suportes instrucionais e monitoramento do progresso do aluno (Achieve, 2016). Infelizmente, a rubrica é consideravelmente extensiva, tornando-a desafiadora de usar, especialmente para professores (Fulmer *et al.*, 2018). Assim, a ferramenta desenvolvida pelo A193 pode ajudar na avaliação e verificação do alinhamento entre a avaliação formativa elaborada e implementada e a instrução.

A segunda questão de pesquisa examinou as práticas de avaliação consideradas como “melhores” pelos professores e as práticas que eles utilizaram em suas avaliações formativas. Verificou-se que os professores geralmente seguiam práticas de articulação de tarefas e inclusão de vários níveis representacionais, entretanto, apenas um dos professores realizou a integração de vários níveis conceituais. Também se notou que os professores conseguiram articular as tarefas, porém com alguma dificuldade. Além disso, eles também apresentaram dificuldades para incorporar diferentes níveis representacionais, o que limitou a eficácia das avaliações na medida da compreensão dos alunos.

A respeito do alinhamento entre as avaliações e os materiais instrucionais, os professores geralmente alinhavam algumas tarefas, mas frequentemente desalinhavam os níveis representacionais. Essa característica se mostrou presente até no professor mais experiente analisado.

A respeito das conclusões do artigo, discute-se que há ferramentas e literatura disponível para melhorar as práticas de avaliação de professores, entretanto a aplicação pode ser difícil. Isso pode ocorrer devido aos diferentes ambientes e critérios de aprendizagem existentes no contexto escolar. Por esse motivo, os professores necessitam de mais do que ferramentas de avaliação; eles precisam de suporte para compreender e aplicar as ferramentas disponíveis de forma eficaz para melhorar suas avaliações e instruções e o alinhamento entre esses.

Similarmente ao artigo A193, muitos outros da categoria C08 avaliaram atitudes de professores, em especial suas práticas avaliativas ou práticas

de avaliação, como anteriormente mencionado. Os artigos pertencentes a esse grupo avaliaram: as melhores práticas para avaliações do tipo inquiry relatadas por professores; as práticas de avaliação e os critérios de avaliação de futuros professores na adequação de vários procedimentos científicos; o uso da avaliação para orientar o ensino; os processos de geração de itens de avaliação de professores; a elaboração de avaliações formativas pelos professores; práticas de avaliações em avaliações somativas; as práticas avaliativas consideradas “melhores” pelos professores.

Trazemos abaixo um excerto do A063 para melhor explicitarmos a natureza dos artigos alocados a categoria C08:

Neste estudo quantitativo, uma versão abreviada do Attitude toward the Subject of Chemistry Inventory (ASCI) criado por Bauer (2008) foi usada para identificar as atitudes químicas de duas populações de estudantes de química geral em duas universidades. O ASCIv2 continha apenas dois fatores do instrumento original. Esses fatores mediam as atitudes emocionais e intelectuais dos alunos em relação à química como disciplina. Os alunos precisaram de apenas dois minutos para responder ao ASCIv2, fornecendo um mecanismo para usá-lo com instrumentos adicionais em um design de pesquisa, ao mesmo tempo em que minimizava a fadiga da avaliação (A063, p. 271, tradução nossa)⁸⁹.

Nota-se que o A063 avalia as atitudes em relação à Química de estudantes no nível superior. De acordo com o A063, embora existam muitos instrumentos para avaliar o conhecimento cognitivo, poucos se concentram em variáveis afetivas como atitudes. A avaliação de atitudes é importante, já que esta pode fornecer informações diagnósticas valiosas relacionadas à metacognição, habilidades de resolução de problemas e crenças dos alunos sobre química, por exemplo.

Para avaliar as atitudes dos alunos, o A063 utiliza uma ferramenta adaptada de Xu e Lewis (2011), o *Attitude toward the Subject of Chemistry Inventory* (ASCIv2). Um dos objetivos principais do A063 é medir as mudanças de atitude nos alunos em cursos de aprendizagem ativa por meio dessa ferramenta. Essa

⁸⁹ No texto original: “In this quantitative study, a shortened version of the Attitude toward the Subject of Chemistry Inventory (ASCI) created by Bauer (2008) was used to identify the chemistry attitudes of two populations of general chemistry students at two universities. The ASCIv2 contained just two factors from the original instrument. These factors measured students’ emotional and intellectual attitudes towards chemistry as a subject. Students required only two minutes to respond to ASCIv2, providing a mechanism to use it with additional instruments in a research design while minimizing assessment fatigue” (A063, p. 271).

ferramenta, originalmente elaborada por Bauer (2008) e posteriormente abreviada pelos autores supracitados, utiliza uma escala diferencial semântica para os alunos classificarem seus sentimentos sobre química. A ferramenta é construída em cima de dois fatores: Acessibilidade Intelectual e Satisfação Emocional e contém adjetivos polares para os alunos auto declararem suas atitudes em relação à Química como uma disciplina. Os adjetivos polares do fator Acessibilidade Intelectual do ASCIv2 foram: Difícil/Fácil; Complicado/Simples; Confuso/Claro; Desafiador/Desafiador. Os adjetivos polares do fator Satisfação Emocional foram: Desconfortável/Confortável; Frustrante/Satisfatório; Desagradável/Agradável; Caótico/Organizado.

O ASCIv2 foi realizado por estudantes em duas universidades dos EUA, sendo que 148 alunos pertenciam a uma universidade do sudeste e 87 de uma universidade do centro-oeste. Os alunos completaram ambas as versões do ASCI em discussões baseadas em *Process-Oriented Guided Inquiry Learning* (POGIL)⁹⁰. Os dados incluíram notas de cursos, pontuações de exames ACS e o ASCIv2, das quais o A063 realiza uma análise quantitativa e estatística.

Em relação aos resultados, o A063 discute que o ASCIv2 detectou mudanças nas atitudes dos alunos em cursos usando pedagogias de aprendizagem ativa de modo efetivo. As análises de correlação feitas pelo A063 mostraram fortes relações entre as pontuações pós-teste do ASCIv2 e o sucesso acadêmico (notas do curso e resultados do exame ACS), indicando uma relação entre sucesso acadêmico e atitudes positivas em relação a Química.

Também se verificou um aumento significativo no fator Acessibilidade Intelectual, sugerindo que as sessões POGIL tornaram a química mais acessível para os alunos. Embora as mudanças na atitude tenham sido pequenas no contexto da universidade do sudeste, o ASCIv2 demonstrou sensibilidade para detectar mudanças nas atitudes em um curto período. Para o A063 isso é relevante, pois reflete o impacto de ambientes de aprendizagem ativa nas percepções dos alunos sobre química. O fator Satisfação Emocional não mostrou mudanças significativas, indicando espaço para maior exploração em pesquisas futuras.

⁹⁰ POGIL, ou Process-Oriented Guided Inquiry Learning, é uma abordagem instrucional que enfatiza o aprendizado ativo por meio de investigação guiada. No POGIL, os alunos trabalham em pequenos grupos em atividades cuidadosamente projetadas que promovem o pensamento crítico e as habilidades de resolução de problemas. Em vez de receber respostas diretas dos instrutores, os alunos exploram dados, modelos e conceitos, assumindo a responsabilidade por seu próprio aprendizado. Este método incentiva a colaboração, a comunicação e a compreensão mais profunda do assunto (A063).

4.2.9 Avalia outras pesquisas (C09)

A categoria C09 consistiu dos artigos que avaliaram pesquisa. No Quadro 29 apresentamos todas as subcategorias encontradas na categoria C08 e os respectivos códigos desses artigos. As subcategorias mostram quais pesquisas foram avaliadas exatamente.

Quadro 29 - C09: Avalia outras pesquisas

Subcategoria	Código
Avalia artigos científicos e a autoavaliação de alunos utilizando o caso da pandemia de COVID-19	A088
Avalia as avaliações usadas em artigos que relataram investigações quantitativas do efeito de intervenções instrucionais em resultados baseados em desempenho	A122
Avalia e analisa a relação entre a pesquisa docente e a eficácia do ensino no doutorado	A143

Fonte: os autores

Trazemos abaixo um excerto do artigo A122 para exemplificar os artigos que avaliaram pesquisa, da categoria C09.

Quase-experimentos são comuns em estudos que estimam o efeito de intervenções instrucionais em resultados de desempenho dos alunos. Nesse tipo de pesquisa, a natureza do desenho experimental, a escolha da avaliação, a seleção de grupos de comparação e os métodos estatísticos usados para analisar os dados de comparação ditam a validade das inferências causais. Portanto, reunir e relatar evidências de validade em estudos causais é de extrema importância, especialmente quando as conclusões têm implicações políticas reais para alunos e professores, entre outras partes interessadas. Esta revisão examina 24 artigos que relataram investigações quantitativas do efeito de intervenções instrucionais em resultados baseados em desempenho conduzidos em cursos de química de graduação (A122, p.401, tradução nossa)⁹¹.

Assim, nota-se que o artigo A122 avaliou pesquisas, especificamente, as avaliações usadas em artigos que realizaram intervenções instrucionais em forma

⁹¹ No texto original: “Quasi-experiments are common in studies that estimate the effect of instructional interventions on student performance outcomes. In this type of research, the nature of the experimental design, the choice in assessment, the selection of comparison groups, and the statistical methods used to analyze the comparison data dictate the validity of causal inferences. Therefore, gathering and reporting validity evidence in causal studies is of utmost importance, especially when conclusions have real policy implications for students and faculty, among other stakeholders. This review examines 24 articles that reported quantitative investigations of the effect of instructional interventions on performance-based outcomes conducted within undergraduate chemistry courses.” (A122, p. 401).

de quase-experimentos. Para o A122, um quase-experimento é um estudo intervencionista empírico usado para estimar o impacto causal de uma intervenção em uma população-alvo sem atribuição aleatória. A pesquisa quase-experimental possui semelhanças com a pesquisa experimental tradicional, mas especificamente carece do elemento de atribuição aleatória para tratamento ou controle. Nas pesquisas quase-experimentais, normalmente permite-se que o pesquisador controle a atribuição para a condição de tratamento.

Em relação ao que foi examinado nos artigos analisados pelo A122, quatro aspectos foram mencionados:

(1) O tipo de desenho quase-experimental usado para estudar a relação entre intervenções, alunos, resultados e cenários; (2) as métricas usadas para medir os resultados de desempenho; (3) o tipo de grupos usados para contrastar os resultados entre as condições experimentais; e (4) os métodos estatísticos usados para analisar os dados de comparação. Por meio do exame desses quatro aspectos dos estudos causais, juntamente com uma tipologia de validade para desenhos quase-experimentais, catalogamos as métricas e métodos usados para comparar os resultados de desempenho dos alunos em vários contextos instrucionais (A122, p.401, tradução nossa)⁹².

Para realizar as análises o A122 analisa artigos que realizaram estudos de intervenção publicados na revista *Chemistry Education Research* publicados entre 2011-2017. Os critérios para seleção dos artigos foi: relatar comparações entre grupos ou dentro do grupo, medir resultados de desempenho, fornecer resultados quantitativos e ser definidos em salas de aula de química no Ensino Superior. Os artigos foram codificados usando um protocolo de leitura e na sequência os autores se reuniram para discussões a fim de verificar a concordância nas suas codificações. Os códigos foram então organizados em matrizes de dados, mostrando a prevalência de cada código por questão de pesquisa.

A respeito dos resultados, o A122 identificou cinco principais métricas utilizadas nas pesquisas para avaliar o desempenho dos alunos e o impacto das intervenções. Esses foram: exames/provas criados pelos próprios professores; taxas

⁹² No texto original: “(1) The type of quasi-experimental design used to study the relationship between interventions, students, outcomes, and settings; (2) the metrics used to measure performance outcomes; (3) the type of groups used to contrast outcomes across experimental conditions; and (4) the statistical methods used to analyze the comparison data. Through the examination of these four aspects of causal studies, together with a validity typology for quasi-experimental designs, we catalogued the metrics and methods used to compare student performance outcomes across varied instructional contexts” (A122, p. 401).

de persistência/retenção; exames/provas do American Chemical Society (ACS); notas; e outras métricas. As métricas mais utilizadas foram os exames/provas criados pelos próprios professores e as taxas de persistência/retenção.

Por fim o A122 discute que nenhum estudo é capaz de cobrir todos os vieses, mas faz algumas recomendações para os pesquisadores. Nesse sentido, esses devem alocar mais atenção a validade, especialmente na priorização do tratamento de ameaças de validade e na apresentação de limitações para orientar pesquisas futuras. Além disso, recomenda-se que futuros estudos que relatam o impacto de intervenções em contextos educacionais em Ensino de Química, alinhem melhor a teoria com a medição/avaliação do desempenho estudantil/impacto da intervenção.

4.2.10 Discute avaliação em contextos diferentes (C10)

A categoria C10 consistiu dos artigos que discutiram a avaliação em um contexto específico. No Quadro 30 apresentamos todas as subcategorias encontradas na categoria C10 e os respectivos códigos desses artigos. As subcategorias mostram quais contextos foram discutidos exatamente.

Quadro 30 - C10: Discute avaliação em contextos diferentes

Subcategoria	Código
Discute a avaliação em uma proposta de ensino	
Discute a avaliação no ensino da Natureza da Ciência	A003
Discute as implicações na avaliação de um ensino e aprendizagem de química não baseado em evidências.	A104
Discute a avaliação na perspectiva do aluno	
Discute a avaliação em diferentes grupos sociais de alunos e a realização de inferências a respeito deles.	A007
Discute e constrói uma estrutura teórica para a avaliação do conhecimento do conteúdo dos alunos	A025
Descreve um processo que mostra como as universidades e outras instituições podem desenvolver e usar medidas de resultados para avaliar a aprendizagem dos alunos.	A050
Discute as observações que fizeram ao analisar o significado de “sucesso” em três ambientes de química geral e a relação com o que foi avaliado.	A196
Discute uma ferramenta de avaliação	
Discute e apresenta um framework para avaliar o raciocínio científico	A019
Discute as potencialidades de avaliações ordenadas de múltipla escolha (OMC)	A093
Discute como novas ferramentas de avaliação podem ser combinadas.	A075
Discute a avaliação na perspectiva do professor	

Discute os impactos negativos da avaliação competitiva sob um ponto de vista pessoal, baseado nas experiências do autor	A052
Discute as mudanças nas práticas de avaliação dos professores, ao serem questionados sobre suas jornadas profissionais e pessoais.	A131
Discute como o arredondamento de notas pode exacerbar as desigualdades sistêmicas e fornece críticas a respeito dessa prática	A204
Discute como as práticas de avaliação no ensino superior podem melhorar ou dificultar a aprendizagem. Apresenta os princípios do Berkeley Evaluation & Assessment Research Assessment System (BEAR) para melhorar o aprendizado.	A199
Discute a avaliação em um currículo específico	
Discute e descreve o sistema de avaliação e o sistema instrucional da disciplina de Ciências Naturais na cidade de Batu	A108
Discute como o programa OCIA!, foi desenvolvido. Discute as avaliações, formativas e somativas empregadas no programa	A127
Discute a importância das avaliações sistemáticas dos programas dos cursos.	A166
Discute dados de avaliação em Química Orgânica que contribuem para a discussão atual a respeito da avaliação adequada nesta disciplina	A110
Discute a necessidade de inovações curriculares e o papel da avaliação	A128
Discute e critica avaliações equivocadas que definem o brilhantismo científico	A045
Discute a avaliação em um contexto tecnológico	
Discute e descreve alguns dos desafios enfrentados na avaliação online e algumas estratégias para superá-los	A151
Discute como a avaliação foi impactada no Ensino de Química online na pandemia	A156
Discute o uso do IMMEX software para a avaliação	A008
Discute e descreve a prática de ensino de aprendizagem online com base em cursos online para professores de química em formação, incluindo as abordagens de avaliação usadas	A167
Discute a avaliação de uma habilidade	
Discute os métodos/ferramentas que tem sido usados para avaliar as habilidades de pensamento sistêmico	A198
Discute as implicações para o pensamento sistêmico na formulação de avaliações	A098
Discute como o desenvolvimento do raciocínio científico e a argumentação pode ser apoiado em avaliações	A142
Discute os esforços de avaliação direcionados para promover e avaliar habilidades espaciais	A140
Discute a avaliação de habilidades cognitivas de ordem superior em Química Ambiental	A203
Discute a avaliação na pesquisa	
Discute o desenvolvimento da pesquisa no campo do ensino de química, com um foco especial nas pesquisas envolvendo avaliação, na universidade de Novi Sad desde a dissolução da República Socialista Federal da Iugoslávia	A161

Fonte: os autores

Ao analisar o Quadro 30, nota-se que a categoria C10 foi uma das mais representativas do *corpus*. A categoria C10 consistiu em 14,1% do *corpus* total (29 artigos), ficando atrás das categorias C03 (Avalia propostas de ensino) e C04 (Avalia um instrumento, ferramenta ou tecnologia), ambas consistindo em 19% e 17,5% do *corpus*, respectivamente. Portanto, discutir avaliação em um contexto específico foi uma tendência de pesquisa nos artigos sobre Avaliação na área de Ensino de Química.

Para exemplificar os artigos alocados na categoria C10, trazemos um excerto do A156 abaixo:

A pandemia da doença do Coronavírus 2019 (COVID-19) na China mudou o ensino superior drasticamente, com um aumento distinto na instrução online. Aqui, descrevemos os principais sucessos e desafios encontrados durante a implementação completa da instrução de química online no ensino superior na China com base em questionários respondidos por 56 professores e 432 alunos em duas universidades. Além de descrever os impactos da instrução de química online em professores, alunos, experimentos de química, avaliações de alunos e tecnologia, nosso trabalho busca fornecer insights sobre como professores, alunos e tecnologias se elevam para enfrentar os desafios neste momento difícil (A156, p. 2810, tradução nossa)⁹³.

Observa-se que o A156 buscou discutir a avaliação em um contexto específico, nesse caso, como a avaliação foi impactada no Ensino de Química online durante o período da Covid, dentre outros pontos como o impacto nos professores, alunos e experimentos. Para isso o A156 utilizou questionários separados para professores e alunos, dos quais tiveram como objetivo identificar suas opiniões sobre o ensino e a aprendizagem online, o impacto da educação online durante a pandemia da COVID-19 e a satisfação com a tecnologia de educação online. Os questionários abertos possuíam escalas de classificação específicas (excelente, muito bom, bom, médio, abaixo da média). As pesquisas foram concluídas por 432 alunos do Ensino Superior e 56 professores, todos de duas universidades encontradas na China. Os

⁹³ No texto original: "The Coronavirus disease 2019 (COVID-19) pandemic in China has changed higher education dramatically, with a distinctive rise in online instruction. Here, we describe the major successes and challenges encountered during the full implementation of online chemistry instruction in higher education in China based on questionnaires answered by 56 teachers and 432 students in two universities. In addition to describing the impacts of online chemistry instruction on teachers, students, chemistry experiments, student assessments, and technology, our work seeks to provide insights into how teachers, students, and technologies rise to meet the challenges in this difficult time" (A156, p. 2810).

alunos eram dos seguintes cursos: Química, Engenharia Química, Educação em Química ou Química Aplicada e Materiais.

A respeito dos resultados do A156, no ensino online, professores e alunos não podiam se ver pessoalmente e se comunicar. Isso é revelado como uma das maiores diferenças entre o ensino online e o presencial. A maioria dos entrevistados (90%) acredita que esse foi o maior inconveniente causado pelo ensino online. Devido a essa característica, houve interação limitada entre os professores e os alunos, algo que foi mencionado por 83% dos respondentes. Além disso, o feedback sobre ensino e aprendizagem foi considerado pequeno e limitado por 73% dos respondentes.

Em relação à avaliação realizada no Ensino de Química online na pandemia, o A156 apresenta alguns desafios. Houve desafios para realizar as avaliações/exames online com sucesso nas disciplinas obrigatórias, estes muitas vezes relacionados à tecnologia. Praticamente metade dos respondentes afirmam que se sentem apenas “bem” ou “não estão satisfeitos” com as avaliações/exames realizadas online. Isso, pois a imparcialidade dos exames/avaliações online é uma questão. Além disso, os alunos precisam usar dois dispositivos eletrônicos nas avaliações, um para concluir o exame online e outro para monitorar os próprios alunos. Exemplos desse monitoramento são dados no A156, como a utilização da webcâmera do aluno apontado para si mesmo, revelando sua bancada de trabalho e consequentemente impedindo a consulta em outros materiais durante a realização da avaliação. Por fim, o A156 faz recomendações para avaliação no período como: o aumento da proporção de avaliações baseadas em desempenho dos alunos e a evitação de testes online sem consulta.

Similarmente ao A156, os outros artigos da C10 discutem avaliação em um contexto específico. No caso do A156 esse contexto é tecnológico, devido ao interesse em compreender o impacto do Ensino de Química online na avaliação, dentre outros pontos. No entanto, outros contextos foram discutidos na C10, como: a avaliação em uma proposta de ensino; a avaliação na perspectiva do aluno; uma ferramenta de avaliação; a avaliação na perspectiva do professor; a avaliação em um currículo específico; a avaliação em um contexto tecnológico; a avaliação de uma habilidade; e a avaliação na pesquisa. Reconhecemos que a C10 se distingue das outras nove categorias, pois nela não há algo se quer sendo avaliado. Os artigos da C10 apenas discutiram avaliação. Assim, a natureza dos artigos da C10 se distingue

dos artigos do restante do *corpus*, pois os artigos da C10 muitas vezes possuem uma natureza teórica e não tão empírica como os artigos encontrados nas categorias C01-C09.

A fim de continuarmos nossas discussões sobre os artigos pertencentes a categoria C10, trazemos um excerto do A204 abaixo:

Os apelos por diversidade, inclusão e equidade na educação em química e ciências têm sido contínuos. No entanto, alguns professores podem achar difícil desafiar as desigualdades sistêmicas devido à sua difusão. Postulamos que uma área de trabalho para abordar as desigualdades sistêmicas na educação em química e ciências é a remoção do arredondamento de notas de nossas práticas. Embora essa prática possa ser vista como tradicional e "o que as pessoas sempre fizeram", oferecemos nove críticas a essa abordagem e como ela pode exacerbar as desigualdades sistêmicas em vez de ajudar os alunos (A204, p. 185, tradução nossa)⁹⁴.

Observamos que o A204 discute avaliação em contextos diferentes. Nesse caso, é discutido a avaliação na perspectiva do professor, mais especificamente, porque a prática avaliativa de arredondar notas é desaconselhado.

As discussões do A204 fornecem um panorama breve da evolução da pesquisa em Ensino de Química. Nele, é destacado uma mudança que ocorreu na área, na qual saímos da dependência das experiências pessoais dos professores para o uso de práticas de ensino e práticas avaliativas baseadas em evidências. Apesar disso, ainda existem práticas educacionais que impactam, negativamente, a equidade e a inclusão principalmente de alunos marginalizados. O A204 enfatiza a importância de examinar criticamente essas práticas ainda fortemente enraizadas na comunidade docente.

As discussões do A204 apontam que os métodos tradicionais de classificação/concessão de nota, especialmente o arredondamento de notas, contribuem para as desigualdades sistêmicas. Isso ocorre devido a uma priorização a classificação e a competição em vez do aprendizado real. Nesse sentido, faz-se necessário uma reformulação completa das práticas de classificação, na qual

⁹⁴ No texto original: "Calls for diversity, inclusion, and equity in chemistry and science education have been ongoing. However, some faculty may find it difficult to challenge systemic inequities due to their pervasiveness. We posit that one working area for addressing systemic inequities in chemistry and science education is the removal of grading on a curve from our practices. Though this practice may be seen as traditional and "what people have always done", we offer nine critiques of this approach and how it may exacerbate systemic inequities rather than help students" (A204, p. 185).

afastamo-nos de métodos que prejudicam os alunos, particularmente aqueles de origens marginalizadas, e adotamos estratégias de avaliação mais favoráveis e inclusivas. Após essas discussões iniciais, o A204 apresenta nove críticas a essa prática avaliativa. Apresentamos alguns dessas críticas na sequência.

Uma das críticas do A204 enfatiza que o arredondamento de notas depende de definições subjetivas do que consideramos como um aluno “na média” ou “normal”. De acordo com A204, isso é problemático, pois esses padrões não levam em consideração a diversidade de origens das quais os alunos vieram e suas múltiplas experiências. Conforme o A204, essa prática afeta especialmente alunos grupos marginalizados. Nesse sentido, questões críticas devem ser levantadas pelos professores a respeito de quem é definido como “normal” e as implicações que isso carrega para equidade no Ensino de Química.

Outra crítica é que o arredondamento de notas contradiz a pedagogia culturalmente relevante. Essa por sua vez enfatiza a necessidade de considerar as diversas origens culturais e necessidades de aprendizagem dos alunos. Essa perspectiva se alinha com a crença de que todos os alunos podem ter sucesso, contrariando diretamente a suposição de que alguns alunos devem inevitavelmente falhar, ou seja, encontrarem-se “abaixo da média”.

Outra crítica apresentada pelo A204 descreve o arredondamento de notas como uma abordagem de “processamento em massa”. Nesse olhar, essa prática prioriza a eficiência em vez da avaliação significativa da aprendizagem do aluno. Logo, esse método corre o risco de tratar os alunos como meros pontos de dados em vez de indivíduos com aprendizagens únicas.

Uma crítica final observa que o arredondamento de notas pode promover uma atmosfera desnecessariamente competitiva entre os alunos, desencorajando a colaboração e inibindo o aprendizado. Nesse olhar, um ambiente educacional competitivo pode favorecer, desproporcionalmente, alguns alunos de origens privilegiadas, que podem estar mais bem preparados para os desafios propostos pelo professor.

O A204 faz um apelo por alternativas ao arredondamento de notas e enfatiza a necessidade de práticas educacionais equitativas em Química. É necessário que os professores se concentram em redesenhar e repensar suas disciplinas para dar suporte a todos os alunos, em vez de usar seus cursos como meras oportunidades preparatórias.

A respeito de alternativas, por mais que isso não seja o objetivo do A204, algumas recomendações são brevemente mencionadas, como: O Design Universal para Aprendizagem (UDL)⁹⁵ e o Design Reverso⁹⁶. De acordo com o A204, mesmo que não seja possível a realização de uma reestruturação completa das disciplinas, a reflexão sobre as práticas de avaliação e classificação já pode gerar melhorias e favorecer a criação de um ambiente educacional mais acessível para todos os alunos.

4.3 QUAIS ARTIGOS TÊM O FOCO PRINCIPAL EM AVALIAR/AVALIAÇÃO?

Além de identificarmos as categorias que evidenciassem o que foi avaliado nos artigos sobre Avaliação na área de Ensino de Química, buscamos identificar a quantidade de artigos do *corpus* que apresentou a avaliação como foco principal na sua estruturação. Dessa maneira poderíamos identificar a quantidade relativa ao *corpus* de artigos que traziam avaliação como um tema central de suas pesquisas.

Para isso, analisamos quais artigos apresentavam o termo avaliação ou seus derivados (avaliar, avaliando, avaliou) nos seus objetivos e/ou perguntas de pesquisa. Apresentamos os resultados desse movimento no Quadro 31.

Quadro 31 – Artigos com foco principal em avaliar/avaliação

Grupo	Códigos
Possui avaliação como um dos seus focos principais	A001, A002, A004, A005, A006, A007, A008, A010, A011, A012, A014, A018, A019, A020, A021, A022, A023, A024, A025, A026, A027, A032, A033, A034, A035, A037, A038, A039, A041, A042, A044, A045, A046, A048, A053, A054, A055, A056, A057, A058, A059, A060, A061, A063, A064, A065, A066, A067, A073, A072, A074, A075, A076, A077, A078, A079, A080, A081, A082, A083, A086, A087, A088, A089, A090, A092, A093, A096, A098, A099, A100, A101, A102, A103, A105, A106, A108, A109, A111, A112, A113, A115, A117, A118, A120, A122, A123, A124, A125, A126, A130, A133, A134, A137, A138, A139, A141, A144, A146, A147, A148, A149, A151, A152, A153, A158, A159, A160, A162, A165, A168, A169, A174, A175, A178, A179, A180, A181, A182, A183, A185, A186, A188, A189, A190, A191,

⁹⁵ Uma estrutura de ensino voltada para criar ambientes inclusivos, removendo barreiras à educação.

⁹⁶ Um framework que alinha as metas de aprendizagem com materiais instrucionais e avaliações, garantindo feedback significativo sobre o domínio do aluno.

	A192, A193, A195, A196, A197, A198, A199, A200, A201, A203, A204, A205
Não possui avaliação como um dos seus focos principais	A003, A009, A013, A015, A016, A017, A028, A029, A030, A031, A036, A040, A043, A047, A049, A050, A051, A052, A062, A068, A069, A070, A071, A084, A085, A091, A094, A095, A097, A104, A107, A110, A114, A116, A119, A121, A127, A128, A129, A131, A132, A135, A136, A140, A142, A143, A145, A150, A154, A155, A156, A157, A161, A163, A164, A166, A167, A170, A171, A172, A173, A176, A177, A184, A187, A194, A202

Fonte: os autores

Nossa expectativa é que os resultados apresentados no Quadro 31 possam ser úteis para outros pesquisadores interessados em pesquisas sobre Avaliação na área de Ensino de Química. Isso porque o Quadro 31 identifica os artigos que possuem um maior enfoque em Avaliação, permitindo a outros pesquisadores um olhar mais direcionado dos artigos que realmente trazem a Avaliação como eixo central do estudo

A respeito do segundo grupo, que inclui os artigos em que a avaliação não é um dos focos principais, consideramos este como importante por alguns motivos. Primeiramente, o segundo grupo oferece um panorama mais amplo sobre como a avaliação está sendo abordada pela área de Ensino de Química de uma maneira mais tangencial ou secundária, também fornecendo a quantidade de artigos em relação ao *corpus* total em que isso ocorreu. Um segundo motivo é que os artigos do segundo grupo podem revelar práticas avaliativas implícitas em contextos nos quais a avaliação não é o eixo central, mas ainda assim exerce um papel relevante nas estratégias pedagógicas e nos processos de ensino e de aprendizagem, já que os próprios autores consideraram relevante o uso da expressão.

Uma terceira razão é que o segundo grupo revela como a avaliação está sendo integrada de uma maneira mais sutil ou complementar no Ensino de Química, podendo trazer insights valiosos para pesquisadores que desejam expandir ou aprimorar o uso da avaliação. A quarta razão é que a segunda categoria pode evidenciar lacunas na literatura, onde a avaliação aparece de forma menos explícita ou meramente complementar. Ao compreender esse uso mais discreto, os pesquisadores interessados em desenvolver pesquisas no tema Avaliação na área de

Ensino de Química novas oportunidades de investigação e até mesmo sugerir maneiras de fortalecer o papel da avaliação nesses contextos.

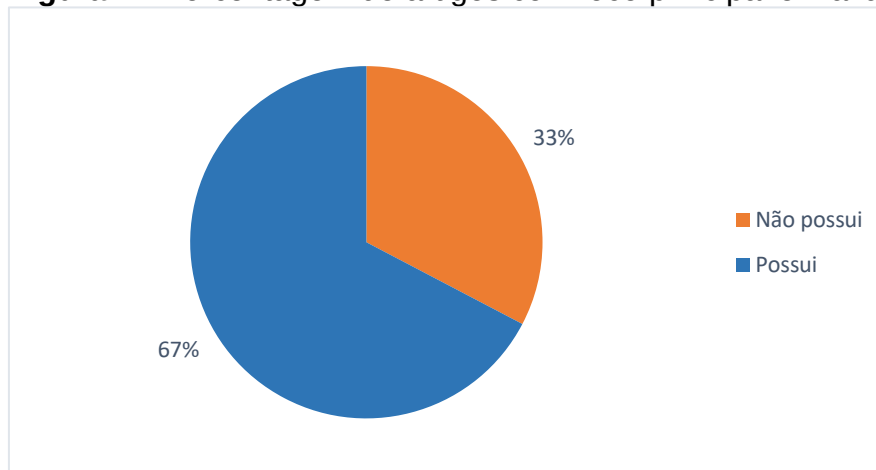
Na sequência apresentamos a quantidade e a porcentagem de artigos com foco principal em avaliação/avaliar do *corpus* no Quadro 32 e na Figura 4.

Quadro 32 – Quantidade de artigos com foco principal em avaliar/avaliação

Grupo	Número de artigos	Porcentagem
Possui avaliação como um dos seus focos principais	138	67%
Não possui avaliação como um dos seus focos principais	67	23%

Fonte: os autores

Figura 4 - Porcentagem de artigos com foco principal em avaliação/avaliar



Fonte: os autores

Apontamos que a maioria dos artigos (67%) trouxe o termo “avaliação” de algum modo em seus objetivos e/ou perguntas de pesquisa e 23% dos artigos não. Logo, há uma porcentagem relevante de artigos sobre Avaliação na área de Ensino de Química que trazem o conceito como um tema central de suas investigações.

Concentramos nossas análises nos artigos do primeiro grupo e identificamos as categorias as quais eles pertencem. Apresentamos esses resultados no Quadro 33.

Quadro 33 – Artigos do primeiro grupo

Categoria	Artigos
C01: Avalia habilidades	A001, A005, A033, A048, A056, A080, A099, A120, A130, A160, A178, A197
C02: Avalia estratégias de avaliação, provas, exames e itens	A002, A021, A026, A039, A042, A058, A077, A086, A090, A092, A113, A118, A124, A134, A146, A195
C03: Avalia propostas de ensino	A004, A020, A046, A067, A072, A073, A081, A082, A100, A102, A103, A105, A106, A109, A149, A152, A165, A169, A191, A192, A205
C04: Avalia um instrumento, ferramenta ou tecnologia	A006, A022, A027, A054, A057, A059, A061, A074, A078, A087, A089, A101, A107, A115, A123, A126, A139, A141, A144, A148, A153, A158, A159, A175, A179, A180, A189, A201,
C05: Avalia compreensões, percepções e interesse	A011, A018, A032, A034, A038, A185, A186
C06: Avalia conhecimentos	A010, A012, A023, A024,
C07: Avalia o currículo	A014, A174, A200
C08: Avalia atitudes	A190, A193
C09: Avalia outras pesquisas	-
C10: Discute avaliação em contextos diferentes	A007, A008, A019, A025, A050, A075, A093, A108, A151, A196, A199, A203, A204

Fonte: os autores

Para ilustrar os artigos alocados no primeiro grupo, artigos que possuem avaliação como um dos seus focos principais, trazemos exemplos de perguntas de pesquisa e ou objetivos de alguns desses artigos na sequência.

Este estudo tem como objetivo produzir: (1) um modelo de pensamento crítico integrado ao processo científico, (2) instrumentos de validade de construto e confiabilidade que **avaliam** as habilidades integradas de pensamento crítico das habilidades do processo científico na aprendizagem de química no ensino médio (A001, p. 285, grifo nosso, tradução nossa)⁹⁷.

Este estudo teve como objetivo investigar a utilização de um método de **avaliação** por pares para melhorar o desempenho de alunos do ensino médio em cálculos de química (A004, p. 569, grifo nosso, tradução nossa)⁹⁸.

Os objetivos deste artigo – Parte II são (a) investigar a aplicação da estrutura de habilidades de transferência de três atributos por meio da realização de dois estudos; e (b) demonstrar o valor da estrutura como

⁹⁷ No texto original: "This study aims to produce: (1) a critical thinking model integrated with the science process, (2) construct validity and reliability instruments that measure the integrated critical thinking skills of science process skills in high school chemistry learning" (A001, p. 285).

⁹⁸ No texto original: "This study aimed to investigate a method of utilising a peer assessment method to improve high school student performance in chemistry calculations" (A004, p. 569).

uma ferramenta para o design de tarefas e **avaliação** das habilidades de transferência dos alunos (A005, p. 154, grifo nosso, tradução nossa)⁹⁹.

Dois objetivos primários de pesquisa orientaram o desenvolvimento deste estudo: (1) Descrever os desafios observados entre estudantes de química em risco em processos de solução relativos ao conceito de mol e itens de **avaliação** de estequiometria. (2) Comparar e contrastar os desafios que se apresentam entre estudantes de química em risco e sem risco para descrever causas potenciais para o desempenho diferencial observado (A012, p. 572, grifo nosso, tradução nossa)¹⁰⁰.

A partir dos exemplos expostos, fica evidente como a avaliação é um conceito norteador das investigações deste grupo, aparecendo nas perguntas de pesquisa ou objetivos dos artigos. Os artigos desse grupo ainda compartilham algumas semelhanças que refletem a importância da avaliação na área de Ensino de Química. Trazemos algumas das principais características comuns a seguir.

Observamos uma tentativa de integração da avaliação às metodologias de ensino utilizadas/investigadas nos artigos. Esse grupo de artigos comumente abordam a avaliação não apenas como uma ferramenta para aferir o conhecimento, mas como parte integrante dos processos de ensino e aprendizagem. Por exemplo, o artigo A001 menciona a criação de instrumentos para avaliar habilidades integradas, o que sugere uma conexão direta entre avaliação e práticas pedagógicas, ou seja, não meramente o que os alunos sabem ou o que memorizaram, mas o que são capazes de fazer (habilidade) com o conhecimento construído.

Também notamos que vários artigos buscaram desenvolver e/ou aplicar novos instrumentos/ferramentas de avaliação. Observamos que essa característica muitas vezes esteve associada a um desejo de investigar a validade ou confiabilidade de novos instrumentos/ferramentas avaliativos. Isso está ilustrado pelo A005, no qual o artigo propõe uma nova estrutura educacional (estrutura de habilidades de transferência de três atributos) e demonstra seu potencial como uma ferramenta avaliativa. O alto volume de artigos alocados a categoria C04: Avalia um

⁹⁹ No texto original: "The goals of this paper – Part II are (a) to investigate the application of the three-attribute transfer skills framework by conducting two studies; and (b) to demonstrate the value of the framework as a tool for design of assignments and assessment of students' transfer skills" (A005, p. 154)

¹⁰⁰ No texto original: "Two primary research objectives guided the development of this study: (1) Describe the challenges observed amongst at-risk chemistry students in solution processes concerning the mole concept and stoichiometry assessment items. (2) Compare and contrast the challenges that presente amongst not-at-risk and at-risk chemistry students to describe potential causes for the differential performance observed" (A012, p.572)

instrumento, ferramenta ou tecnologia também reverbera esse achado, já que essa categoria foi a segunda mais representativa do *corpus* (17,5% dos artigos). Nesse sentido, acreditamos que a ênfase em novas ferramentas/instrumentos avaliativos dos artigos reflete o desejo por propor abordagens inovadoras para a avaliação na área de Ensino de Química.

A revisão da literatura conduzida por Broietti, Santin Filho e Passos (2015) no contexto nacional apresentou um resultado semelhante, no que diz respeito as ferramentas/instrumentos avaliativos. O estudo supracitado analisou 17 artigos da temática de avaliação publicados em periódicos nacionais entre 1986-2012 e alocou-os em três categorias: o uso de instrumentos avaliativos, a análise de questões e/ou programas de avaliação, e a avaliação e a formação de professores. Na primeira categoria, o uso de instrumentos avaliativos, foram alocados cinco artigos (29%). Nesses artigos, são descritas situações em que, para avaliar um conteúdo específico ou uma determinada competência, utilizou-se um instrumento avaliativo. Os instrumentos mencionados no contexto nacional foram: mapas conceituais, diagramas heurísticos, formulação de questões como instrumentos de avaliação da aprendizagem em Química, para a avaliação de conteúdo ou competências.

Já os artigos pertencentes ao segundo grupo, os que não trouxeram avaliação como tema central em suas perguntas de pesquisa e objetivos, apresentaram uma grande variedade e diversidade de enfoques. Nesses artigos o papel da avaliação aparece como secundário, muitas vezes sendo mencionado apenas ao fim dos resultados e discussões ou na seção de considerações finais, de maneira breve. Apesar disso, discutimos as semelhanças a respeito de como os artigos do segundo grupo tratam a avaliação na sequência.

Observamos uma tentativa de integrar a avaliação com objetivos específicos de aprendizagem presentes nos artigos. A maioria dos artigos pertencentes ao segundo grupo enfatiza, mesmo que de maneira breve e frequentemente sem aprofundamento, que a avaliação deve se alinhar estreitamente com os objetivos de aprendizagem, seja da proposta de ensino desenvolvida/implementada ou do contexto educacional analisado no artigo. Por exemplo, o Artigo A003 destaca uma lacuna na avaliação da natureza da ciência, enquanto o Artigo A016 usa avaliações externas para avaliar a eficácia de uma tarefa de Escrita para Aprender.

Outra semelhança foi a variedade e a diversidade notável de métodos de avaliação dos artigos do segundo grupo. O Artigo A013, por exemplo, empregou questionários e realizou o rastreamento de contagem de acertos para analisar o engajamento dos alunos com um *Maths Skills Site*, enquanto o A015 usou planilhas de avaliação de habilidades de processo científico específicas e ferramentas de observação. Para continuarmos as exemplificações dos vários métodos avaliativos utilizados nos artigos, o A029 discutiu o uso de pré e pós-testes para avaliar os resultados da aprendizagem em um ambiente baseado em jogos. Consideramos que essa variedade indica o reconhecimento de que diferentes contextos e ambientes de aprendizagem exigem estratégias de avaliação distintas para avaliar efetivamente a aprendizagem do aluno, seu engajamento, a qualidade ou potencial de uma intervenção ou proposta de ensino, dentre outros.

Também observamos que alguns artigos alocados ao segundo grupo trouxeram o termo avaliação nas suas conclusões ou considerações finais ao discutirem o uso da avaliação para informar a instrução docente. Nesses casos, a avaliação foi frequentemente retratada como uma ferramenta para informar e melhorar práticas instrucionais, práticas docentes ou práticas de ensino, etc. No Artigo A017, por exemplo, os autores analisaram as respostas dos alunos a avaliações abertas para extrair insights sobre o alinhamento instrucional. Já o Artigo A030 destacou o desenvolvimento de um modelo preditivo baseado em dados de avaliação para adaptar intervenções para alunos despreparados. Isso ilustra como os pesquisadores têm percebido a avaliação como um mecanismo não apenas para avaliar a aprendizagem do aluno, mas também para ajustar/alinhar o ensino do professor.

Finalmente, observamos que alguns artigos apenas discutiram os desafios associados à implementação de práticas de avaliação eficazes para os contextos educacionais específicos investigados. O A015, por exemplo, discute a necessidade de formação em habilidades de processo científico para preparar futuros professores de química, ressaltando uma lacuna de conhecimento e práticas a respeito de avaliação entre educadores. Já o A017 revela dificuldades que os alunos enfrentam ao articular declarações explicativas em avaliações, sugerindo uma desconexão entre as expectativas instrucionais e o desempenho do aluno.

4.4 UMA SÍNTESE CRÍTICA DAS TENDÊNCIAS E LACUNAS - PARTE II

Ao analisarmos o *corpus* de 205 artigos sobre avaliação na área de Ensino de Química, reunimos as tendências mais significativas no que diz respeito ao que mais foi avaliado nos artigos, ou seja, o que a área tem mais valorizado nas pesquisas sobre essa temática. Essas tendências, suas características principais, lacunas e críticas a respeito de sua prevalência são discutidas a seguir.

A categoria que mais se destacou foi a C03: Avalia propostas de ensino. Ela representou 19% do total de artigos. Essa predominância de artigos que avaliam propostas de ensino indica um forte interesse dos pesquisadores em examinar a eficácia, qualidade e impacto de diversas abordagens pedagógicas, intervenções e atividades educativas. As propostas avaliadas variaram amplamente, consistindo desde frameworks até workshops e sequências didáticas. A variedade de propostas de ensino nos sugere que os pesquisadores buscam não apenas implementar novas metodologias, mas também compreender como elas têm influenciado a aprendizagem dos alunos.

Apesar disso, essa tendência de avaliar propostas de ensino, traz consigo algumas limitações em seus resultados. Notamos que muitos desses artigos dificultam a elaboração de comparações e generalizações, devido ao fato de terem sido realizadas em contextos específicos. Dessa forma, os resultados nem sempre podem ser extrapolados em outras práticas ou em outros contextos, embora ofereçam discussões e reflexões a respeito da avaliação de um significativo número de propostas de ensino em contextos variados. Assim, reconhecemos a necessidade de futuras pesquisas que busquem não apenas avaliar propostas de forma individuais, mas que também investiguem aplicabilidade em diferentes contextos educativos.

A segunda tendência percebida foi artigos que avaliam instrumentos, ferramentas ou tecnologias. Isso foi apoiado pela representatividade da categoria C04: Avalia um instrumento, ferramenta ou tecnologia. Essa categoria foi a segunda categoria mais representativa, com 17,5% do *corpus*. Isso indica uma tendência nas pesquisas em avaliar recursos tecnológicos e instrumentos de ensino. Acreditamos que a inclusão de novas tecnologias no ensino de química representa uma tendência crescente e significativa nas pesquisas educativas. Estabelecemos uma relação entre esse achado e a consciência da importância de integrar ferramentas mais contemporâneas nas práticas pedagógicas que os pesquisadores na área de Ensino

de Química têm adquirido nos últimos anos. Essa conscientização está estritamente relacionada a digitalização e o acesso, cada vez mais precoce, dos alunos à tecnologia. Logo, acreditamos que os pesquisadores reconhecem quando os alunos se beneficiam de abordagens que utilizam tecnologias interativas e acessíveis, e estes estão interessados em investigar/avaliar suas potencialidades.

A terceira tendência foi a alta incidência de artigos alocados na categoria C10: Discute avaliação em contextos diferentes, que representou 14,1% do *corpus*. Apontamos que embora esses artigos não realizaram avaliações empíricas diretas de sujeitos ou objetos no contexto educacional, as discussões teóricas aprofundadas traçadas nessas pesquisas são importantes para contextualizar a prática avaliativa. Essa tendência indica que, além de aplicar e avaliar, os pesquisadores estão interessados em reflexões profundas e exaustivas sobre as práticas de avaliação.

A quarta tendência percebida diz respeito ao foco dos artigos (foco em alunos e foco em professores). Observou-se uma tendência em avaliar as percepções de ambos os sujeitos envolvidos nos processos de ensino de aprendizagem, tanto os alunos como os professores. Na categoria C05, que avalia compreensões e percepções, 68% dos artigos focaram nos alunos e 32% direcionaram suas investigações nos professores. Reconhecemos essa característica como uma potencialidade das pesquisas em Avaliação na área de Ensino de Química, pois os pesquisadores estão direcionando suas investigações tanto para os discentes, como para os docentes.

Os artigos que investigaram as percepções dos alunos contribuem para uma compreensão mais profunda de como estes constroem o conhecimento. Acreditamos que essa tendência de focar no aluno representa um esforço para melhorar o processo educativo a partir das necessidades e percepções dos próprios estudantes. O foco crescente nos alunos reflete uma valorização do papel ativo do estudante no processo educativo. No caso da avaliação, essa característica do *corpus* pode gerar melhores práticas de avaliação e de ensino, já que as pesquisas estão sendo baseadas em dados mais concretos sobre o que realmente impacta o aprendizado dos alunos. Assim, ao focar mais nos alunos, as pesquisas em Avaliação na área de Ensino de Química estão refletindo um ensino de maneira mais centrada no estudante.

Em conclusão, as quatro tendências observadas nos artigos sobre avaliação na área de Ensino de Química indicam o que a área tem mais valorizado nesse campo de pesquisa. O enfoque em avaliação de propostas de ensino, ferramentas e tecnologias, e a elaboração de discussões teóricas sobre avaliação, demonstram um interesse em verificar a qualidade e a credibilidade de propostas de ensino, ferramentas de ensino e novas tecnologias educacionais. Além disso, as tendências ressaltam o interesse da área em refletir a respeito do papel, evolução e características da avaliação em contextos diversos.

4.5 UM RETORNO AO TRIÂNGULO DA AVALIAÇÃO

Nesta subseção fazemos um retorno ao triângulo da avaliação, Figura 1, como proposto pelo NRC (2001), constituído pelos vértices cognição, observação e interpretação. No vértice da cognição estão as estruturas mentais e processos mobilizados pelo sujeito avaliado, ou seja, está relacionado com as maneiras pelas quais o sujeito interpreta e organiza as informações, desenvolve habilidades ao resolver uma tarefa. No vértice da observação, incluem-se as tarefas ou atividades propostas, aquilo que fornece a evidência. Por sua vez, no vértice da interpretação estão os critérios de análise a partir dos quais os dados observados serão interpretados. Nesta seção fizemos um exercício de buscar aproximações desse modelo com as categorias evidenciadas nesta pesquisa, em especial as categorias C01 a C08. Dessa forma, esperamos que professores, pesquisadores e formadores possam compreender de que maneira o triângulo apresentado no capítulo inicial pode ser utilizado no contexto avaliativo, para avaliar: habilidades (C01); estratégias de avaliações (C02); propostas de ensino (C03); instrumentos, ferramentas e tecnologias (C04); compreensões (C05); conhecimentos (C06); currículo (C07); e atitudes (C08). Também esperamos que este modelo auxilie na compreensão de como cada um dos vértices do triângulo (cognição, observação e interpretação) pode ser interpretado na avaliação dos elementos supracitados. Optamos por não incluir as categorias C09 e C10 nessa subseção, pois na categoria C09 os autores avaliaram outras pesquisas/artigos e na categoria C10 discutiram a respeito da avaliação em contextos diferentes.

O triângulo da avaliação envolve a observação, a interpretação e a cognição compondo os três vértices, como visto na Figura 1. Ele se caracteriza como

um modelo para compreender os diferentes componentes envolvidos em uma avaliação, e suas relações/interações. O triângulo pode ser utilizado para entender como os sujeitos interpretam e organizam as informações, evidenciando o processo pelo qual educadores coletam e interpretam informações para na sequência elaborarem e realizarem inferências.

4.5.1 HABILIDADES E O TRIÂNGULO DA AVALIAÇÃO

A categoria C01: Avalia habilidades incluiu artigos que abordaram a avaliação de habilidades como pensamento crítico, abstração, raciocínio gráfico, pensamento científico, resolução de problemas, leitura, entre outras. Ao refletirmos sobre a avaliação dessas habilidades, com base nos vértices do triângulo de avaliação, é possível estabelecer algumas associações.

Em relação à observação, é fundamental identificar as habilidades em foco, como, por exemplo, resolução de problemas, pensamento científico e pensamento crítico, que são mencionadas nos artigos da C01. Tomando como exemplo a habilidade de resolução de problemas, é importante observar como o aluno interpreta a tarefa, coleta informações, utiliza estratégias e resolve o problema. A observação pode ocorrer de diferentes formas: diretamente em tarefas ou atividades (como aulas práticas de laboratório, trabalhos em grupo etc.); por meio de materiais produzidos pelos alunos (portfólios, projetos, textos, mapas etc.); ou observando os comportamentos dos alunos (como a comunicação e as estratégias utilizadas para resolver problemas).

No que diz respeito à interpretação na avaliação de habilidades, as observações devem ser analisadas de maneira cuidadosa e a partir de critérios definidos previamente e compartilhados com os sujeitos avaliados. Por exemplo, se um aluno identifica rapidamente um problema, mas tem dificuldade em encontrar uma solução, isso pode ser interpretado como uma limitação na habilidade de resolução de problemas. Além disso, é importante considerar as particularidades do contexto e do aluno para aumentar a confiabilidade da interpretação e do julgamento final. Deve-se levar em conta o conhecimento prévio do aluno sobre os temas relacionados à habilidade em desenvolvimento, sua interação com os colegas, caso a habilidade seja trabalhada em um ambiente de grupo, por exemplo. Também é relevante observar a progressão do aluno ao longo do tempo.

No que diz respeito à cognição na avaliação de habilidades, é necessário definir claramente a habilidade, compreendê-la conceitualmente, identificar seus componentes essenciais e conhecer os indicadores de um elevado nível de domínio. Também é fundamental considerar a complexidade da tarefa ou do contexto em que o aluno demonstra suas habilidades. Por fim, é preciso tomar uma decisão sobre o nível de proficiência alcançado pelo aluno. Por exemplo, o aluno demonstra um alto nível de domínio ao aplicar técnicas avançadas para resolver um problema, ou ainda está em um estágio mais básico, como na identificação do problema ou no levantamento de hipóteses para solucioná-lo?

4.5.2 ESTRATÉGIAS DE AVALIAÇÃO, PROVAS, EXAMES E ITENS E O TRIÂNGULO DA AVALIAÇÃO

Ao refletirmos a respeito da avaliação de estratégias de avaliação, provas, exames e itens e os vértices do triângulo de avaliação, foi possível estabelecer algumas associações.

Acreditamos que o triângulo de avaliação pode ser usado para avaliar a qualidade de avaliações, propriamente ditas. Ao pensarmos nos três componentes: cognição, observação e interpretação, podemos avaliar se uma avaliação se caracteriza como uma ferramenta eficaz para medir a aprendizagem do aluno (avaliação somativa), compreender o conhecimento de um aluno e propor feedback para abordar lacunas e fazer melhorias (avaliação formativa) ou ocasionar a reflexão do aluno e desafiá-lo a confrontar suas limitações e desenvolver sua autorregulação (autoavaliação).

Primeiro, a cognição analisa o que a avaliação mede. Aqui é necessário identificar o que a avaliação de fato é capaz de “capturar”, seja conhecimento, habilidades, percepções, atitudes, etc. Nesse sentido é preciso comparar as expectativas cognitivas e as limitações da estratégia de avaliação, prova, conjunto de itens, etc., verificando se há um alinhamento. Em seguida, a observação cumpre o papel de reunir dados ou evidências do aprendizado do aluno.

Finalmente, a interpretação se concentra na análise das evidências coletadas da avaliação e a formação de um julgamento. Deve-se considerar se os resultados da avaliação podem ser interpretados com um nível de confiabilidade satisfatório. Além disso, pode-se verificar se os critérios de avaliação são claros e consistentes. Aqui, a subjetividade do avaliador também entra em jogo, portanto é

necessário ter critérios ou indicadores que nos auxiliem a considerar diferentes fatores que serão considerados em cada situação avaliativa.

Assim, concluímos que o triângulo da avaliação pode ser aproximado dessa categoria para avaliar se uma avaliação existente é bem projetada e alinhada para medir/compreender o aprendizado ao qual se destina, contemplando avaliações do tipo somativo ou formativa, por exemplo. Acreditamos que possa haver maiores limitações para avaliações do terceiro tipo, autoavaliações, ou seja, aquelas que promovem a autorregulação, conforme Earl (2012). Estudos mais aprofundados que fazem confrontos entre o triângulo e esse tipo de avaliação pode ocasionar discussões e reflexões válidas.

4.5.3 PROPOSTAS DE ENSINO E O TRIÂNGULO DA AVALIAÇÃO

Ao refletirmos a respeito da avaliação de propostas de ensino e os vértices do triângulo de avaliação, foi possível estabelecer algumas associações. O triângulo de avaliação pode ser adaptado e usado para avaliar propostas de ensino (tarefas, intervenções, workshops, estruturas, etc.) concentrando-se nos três elementos principais, representado pelos vértices: cognição, observação e interpretação.

Nesse caso a cognição se refere aos objetivos de aprendizagem que a proposta visa atingir. Para tarefas, isso pode ser as habilidades cognitivas a serem avaliadas (por exemplo, pensamento crítico, resolução de problemas). Para intervenções ou workshops, isso pode estar associado as habilidades que a proposta pretende desenvolver, como melhorar a compreensão conceitual ou o envolvimento do aluno.

A observação envolve coletar evidências de como os alunos se envolvem com a proposta de ensino, podendo ser por meio da verificação das respostas dos alunos às tarefas, a observação de como os professores implementaram estratégias durante a proposta de ensino ou a observação do envolvimento dos participantes nas atividades da proposta de ensino. Assim, dados para avaliar o quão bem a proposta de ensino foi elaborada/executada são gerados.

A interpretação das propostas de ensino implica na análise das evidências observadas para determinar o sucesso da proposta de ensino. Isso envolve avaliar: as potencialidades percebidas da proposta, o engajamento dos

alunos com a proposta, a eficácia da proposta em alcançar os objetivos pretendidos no início, as respostas ou comportamentos dos alunos, etc. Também devem ser considerados, nesse momento, o contexto e os fatores externos que podem influenciar os resultados.

4.5.4 INSTRUMENTOS, FERRAMENTAS E TECNOLOGIAS E O TRIÂNGULO DA AVALIAÇÃO

O triângulo de avaliação pode ser usado para avaliar instrumentos, ferramentas ou tecnologias educacionais focando nos três componentes principais: cognição, observação e interpretação.

Nesse caso, a cognição envolve definir quais habilidades ou conhecimentos a ferramenta, instrumento ou tecnologia visa desenvolver nos alunos. Por exemplo, se a tecnologia é voltada para superar as limitações matemáticas dos alunos ao resolverem problemas na química, a cognição pode envolver habilidades como resolução de problemas, raciocínio lógico e a compreensão de conceitos matemáticos essenciais para a química. Outro exemplo seria a avaliação de mapas conceituais como ferramentas para avaliar. Nesse caso seria necessário definir o que o mapa conceitual pretende desenvolver (seja habilidades, conhecimento, domínio de conceitos) primeiramente. Por exemplo, ele visa ajudar os alunos a organizar os conhecimentos, hierarquizá-los, buscando uma reconciliação integrativa. Avaliar a clareza dos objetivos de aprendizagem da ferramenta, como a organização do conhecimento, é essencial.

No caso da observação é necessário observar como os alunos interagem com a ferramenta ou tecnologia. Isso pode incluir como eles utilizam a plataforma, interagem com o conteúdo, engajam com as estratégias da proposta. Assim, a observação pode ser feita por meio da observação direta, observação do feedback dos alunos ou análise de seu desempenho ao usar a ferramenta/tecnologia. No exemplo de um mapa conceitual, pode ser observado como os alunos interagem com o mapa conceitual, por meio do acompanhamento do processo de construção do mapa e a observação da capacidade dos alunos de identificarem e conectarem conceitos corretamente. A coleta de dados pode incluir análises de mapas conceituais criados pelos alunos, entrevistas sobre como eles estruturaram suas ideias e até feedback dos próprios alunos sobre a ferramenta.

Após a coleta dos dados, é necessário interpretar os resultados para avaliar a eficácia do instrumento ou tecnologia, caracterizando o momento de interpretação, segundo o triângulo. Isso envolve analisar se os estudantes alcançaram os objetivos de aprendizagem pretendidos e se a ferramenta realmente facilita o aprendizado. Aqui, é importante considerar o contexto em que a ferramenta ou tecnologia foi usada, os participantes, sua familiaridade com a tecnologia, sua acessibilidade e a adaptação ao conteúdo de ensino. No exemplo do mapa conceitual como uma ferramenta, isso envolve avaliar se o seu uso ajudou os alunos a alcançarem os objetivos de aprendizagem definidos, se os alunos melhoraram sua compreensão do conteúdo e sua capacidade de fazer conexões entre diferentes conceitos.

Logo ao aplicar o triângulo de avaliação dessa forma, é possível determinar se o instrumento, ferramenta ou tecnologia foi eficaz no suporte ao processo aprendizagem, identificar áreas de melhoria e entender o impacto no desenvolvimento cognitivo dos alunos.

4.5.5 COMPREENSÕES, PERCEPÇÕES E INTERESSE E O TRIÂNGULO DA AVALIAÇÃO

O triângulo da avaliação pode ser uma ferramenta útil para avaliar não somente o conhecimento que os alunos possuem sobre fenômenos, reações e representações químicas, mas também para investigar as suas percepções, compreensões e interesses sobre esses temas.

A cognição nesse caso envolve entender como os alunos estruturam e organizaram seus conhecimentos sobre os fenômenos químicos, reações químicas e estruturas. Os modelos de cognição dos alunos podem incluir representações mentais dos conceitos químicos como a tabela periódica, equações de reações químicas, ou modelos atômicos e como os alunos progridem no entendimento dessas ideias. A avaliação do seu desenvolvimento cognitivo pode ser realizada pelo mapeamento da evolução das ideias dos alunos sobre, por exemplo, o modelo atômico, iniciando em uma visão elementar de uma esfera indivisível até uma compreensão mais sofisticada dos elétrons em orbitais, e até mesmo conceitos mais avançados como os expressos na teoria ácido-base de Lewis

A respeito da avaliação de percepções e Interesses, o vértice cognitivo também pode ser usado por meio de perguntas abertas ou discussões que

revelam as noções dos alunos a respeito de fenômenos, reações e representações (percepções e compreensões) e se os alunos veem a química como algo relevante e interessante (interesse) ou, se por outro lado, os alunos possuem percepções de que os fenômenos químicos são distantes ou difíceis de entender.

A observação no triângulo de avaliação refere-se às atividades que podem ser feitas para provocar as respostas dos alunos e fornecer evidências claras sobre sua compreensão. No contexto de temas químicos, isso pode ser feito no meio de tarefas práticas (experimentos ou simulações) a fim de observar como os alunos compreendem fenômenos químicos. Por exemplo, ao observar a reação entre um ácido e uma base, pode-se analisar se os alunos são capazes de identificar e descrever corretamente as mudanças observadas e conectar essas observações com o conceito de pH. Outra opção são as representações dos alunos (desenhos, textos explicativos, reações propriamente ditas) a respeito de fenômenos químicos, por exemplo. Essas podem incluir equações balanceadas, diagramas de estruturas moleculares ou mecanismos de reação. Assim, é possível observar como os estudantes lidam com a abstração e a aplicação dos conceitos químicos em novos contextos. Por fim, uma outra opção são discussões em grupo ou respostas escritas a questionários.

A interpretação está relacionada à análise das evidências coletadas durante a observação, com o objetivo de tirar conclusões sobre o nível de compreensão dos alunos, suas percepções e interesses. Nesse caso, pode-se realizar a análise das respostas dos alunos, principalmente em suas tentativas de descrever/explicar uma reação química ou o acontecimento de um fenômeno químico. Pode ser considerado a profundidade da compreensão, a integração de conceitos, a utilização de conceitos mais complexos, etc.

Em resumo, o triângulo da avaliação pode ser usado para entender e avaliar as percepções, compreensões e interesses dos alunos em relação a conceitos químicos ao integrar as observações práticas, a análise dos modelos de cognição empregados pelos alunos e a interpretação das evidências coletadas.

4.5.6 CONHECIMENTOS E O TRIÂNGULO DA AVALIAÇÃO

O triângulo da avaliação pode ser utilizado para avaliar o conhecimento dos alunos sobre a química, integrando os três vértices: cognição, observação e interpretação.

No vértice cognitivo, é importante entender como os alunos constroem, interpretam e organizam o conhecimento sobre os conceitos químicos. Isso envolve identificar como eles percebem e relacionam ideias. Esse vértice compreende como os alunos progredem em seus conhecimentos, de noções básicas como “átomos são pequenas partículas” para um entendimento mais complexo, envolvendo elétrons, prótons e interações moleculares.

Nesse caso, os modelos de cognição ajudam a avaliar como eles avançam na compreensão dos conceitos e suas aplicações. Citamos como alguns exemplos de modelos de cognição: esquemas cognitivos, ou seja, representações mentais que os alunos formam sobre conceitos químicos, como átomos e ligações (Bransford *et al.*, 2000); modelos de progressão de aprendizagem, que descrevem como o conhecimento evolui de conceitos simples para mais complexos (NRC, 2001); metáforas ou analogias (Stowe & Cooper, 2019); e modelos de resolução de problemas, que permite que os alunos apliquem o que aprenderam em situações novas e desafiadoras (Bransford *et al.*, 2000)

No vértice da observação, a avaliação ocorre por meio de tarefas específicas que permitem captar a forma como os alunos aplicam o conhecimento. Atividades como a resolução de problemas químicos ou a realização de experimentos práticos, ajudam a observar como os alunos aplicam os conceitos. Também pode ser observada a capacidade dos alunos de representar graficamente esses conceitos, como ao desenhar estruturas de moléculas ou equilibrar reações.

Por fim, no vértice da interpretação, as evidências coletadas por meio da observação são analisadas. Ao avaliar as respostas dos alunos, o professor identifica se os alunos realmente compreenderam os conceitos ou realizaram apenas uma memorização superficial. Aqui, essas evidências podem ser observações diretas, como respostas escritas, demonstrações práticas ou até mesmo discussões orais em sala de aula. A função do professor será interpretar essas evidências para verificar se o aluno conseguiu integrar os conceitos ou se está trabalhando apenas no nível da reprodução.

Assim, o triângulo da avaliação integra cognição, observação e interpretação para fornecer uma visão abrangente do conhecimento dos alunos,

permitindo identificar não apenas o nível de entendimento dos conceitos, mas também as áreas que necessitam de mais atenção.

4.5.7 CURRÍCULO E O TRIÂNGULO DA AVALIAÇÃO

O triângulo da avaliação pode ser utilizado para avaliar e ajustar um currículo, verificando seu alinhamento com os objetivos de aprendizagem e as necessidades dos alunos.

No vértice da cognição, a avaliação do currículo considera como ele favorece o desenvolvimento do conhecimento e/ou habilidades dos alunos. O currículo precisa ser projetado de forma a atender a diferentes níveis de compreensão dos alunos e apoiar a construção de conceitos de maneira progressiva. A análise do currículo por meio do vértice da cognição ajuda a identificar se ele segue uma progressão de aprendizagem lógica e se oferece atividades, projetos e conteúdos que sustentam o desenvolvimento do raciocínio previamente pretendido e uma compreensão mais profunda daquele tema.

O vértice da observação se refere a como o currículo é implementado na prática e como ele é percebido pelos alunos. A avaliação do currículo nesse momento envolve observar como as tarefas propostas geram respostas dos alunos, seja em forma de trabalhos, discussões, testes ou outras atividades. Por exemplo, se o currículo propõe uma sequência de atividades práticas, como experimentos químicos, as respostas dos alunos podem ser observadas para avaliar se eles estão realmente compreendendo os conceitos ou apenas seguindo procedimentos de forma mecânica. A observação ajuda a perceber se o currículo está realmente engajando os alunos e favorecendo a aplicação dos conceitos em contextos reais.

No vértice da interpretação, a análise do currículo envolve a interpretação das evidências coletadas ao longo do processo de avaliação. Isso inclui as respostas dos alunos, os feedbacks dados e as observações feitas nas atividades de aprendizagem. A interpretação ajuda a entender se o currículo está atingindo os objetivos de aprendizagem e se está promovendo uma compreensão profunda dos conteúdos. Caso as evidências mostrem que os alunos não estão conseguindo compreender os conceitos de forma significativa ou que estão apenas repetindo informações sem entendimento real, isso pode indicar que o currículo precisa ser ajustado. Além disso, a interpretação envolve analisar se o currículo atende a

diferentes estilos de aprendizagem e se consegue engajar os alunos, considerando também suas atitudes e interesses em relação ao conteúdo.

Embora o triângulo da avaliação possa ser usado para avaliar um currículo, reconhecemos que ele possui algumas limitações. Primeiramente, o triângulo foca principalmente nos aspectos cognitivos e nas observações, deixando em segundo plano fatores subjetivos, como a motivação dos alunos e influências externas que também impactam o aprendizado. Apesar de existir o terceiro vértice da interpretação, sentimos que a estruturação é limitada para a avaliação curricular. Além disso, reconhecemos que a análise baseada na observação e interpretação das respostas é altamente subjetiva e depende grandemente da percepção pessoal que o avaliador possui. Aqui destacamos a importância de critérios avaliativos previamente definidos e que estes sejam acordados com os estudantes.

Currículos mais dinâmicos com características flexíveis podem dificultar a aplicação consistente do triângulo da avaliação, já que ele depende de dados estruturados e evidências claras para fazer uma análise eficaz. Assim, o triângulo pode não abranger todas as metodologias de ensino inovadoras ou abordagens pedagógicas mais dinâmicas, limitando a avaliação completa de algumas práticas que não se enquadram em um modelo mais tradicional.

Por essas razões, sugerimos o uso do triângulo de forma complementar a outras ferramentas de avaliação, quando se diz respeito a avaliação de um currículo.

4.5.8 ATITUDES E O TRIÂNGULO DA AVALIAÇÃO

No contexto da química, o triângulo da avaliação pode ser uma ferramenta para avaliar atitudes, como as emocionais, intelectuais, positivas e práticas de avaliação. Cada um dos três vértices do triângulo (cognição, observação e interpretação) pode ser utilizado para identificar e compreender como as atitudes de alunos e/ou professores afetam seu aprendizado.

As atitudes dos alunos influenciam diretamente como eles abordam os conceitos químicos. Atitudes emocionais, como o medo de errar/falhar, a ansiedade durante provas/tarefas e a frustração na resolução de problemas difíceis, podem afetar a forma como os alunos aprendem. Similarmente, atitudes intelectuais, como

curiosidade ou desejo de explorar e entender fenômenos do cotidiano em um nível científico, promovem uma aprendizagem mais profunda.

O vértice da cognição permite avaliar como as atitudes emocionais e intelectuais afetam o processo de aprendizagem. Por exemplo, um aluno com uma atitude positiva, como motivação, pode demonstrar maior disposição em resolver problemas desafiadores. Por outro lado, um aluno com uma atitude negativa, como o medo de errar ou desinteresse, pode apresentar dificuldades em aprender e compreender os conceitos. Durante a aprendizagem, o professor pode analisar as atitudes dos alunos e verificar se estão enfrentando os desafios com perseverança e disposição, usando isso como indicador de suas atitudes, por exemplo.

Atitudes como a ansiedade, medo de errar ou desinteresse podem ser observadas em como os alunos participam das aulas e atividades práticas. O professor pode observar o comportamento dos alunos durante experimentos, discussões em grupo ou atividades que requerem maior engajamento dos alunos. Por exemplo, alunos com níveis mais elevados de ansiedade podem evitar fazer perguntas ou hesitar na realização de experimentos laboratoriais. Por outra lado, alunos com atitudes positivas podem se engajar com os desafios de forma proativa. Assim, a observação permite ao professor perceber padrões de comportamento que podem refletir atitudes emocionais e/ou intelectuais dos alunos.

A interpretação envolve analisar se as atitudes estão facilitando ou dificultando o aprendizado. Ao analisar o desempenho dos alunos, o professor pode identificar se um aluno com atitudes negativas está apenas memorizando informações sem compreender os conceitos, ou se está evitando o conteúdo por medo de falhar, por exemplo. Em contrapartida, alunos com atitudes positivas, como perseverança e curiosidade, podem demonstrar um bom entendimento, mesmo em tópicos complexos. A interpretação das respostas, como a maneira como o aluno resolve problemas ou responde a questões durante um teste de química, pode revelar as atitudes subjacentes ao seu desempenho.

Apesar do exposto, reconhecemos que avaliar atitudes usando o triângulo da avaliação apresenta limitações importantes, especialmente nos vértices da interpretação e da observação. A interpretação das atitudes pode ser desafiadora, pois não é possível saber exatamente o que o aluno está pensando ou quais são suas motivações internas. O comportamento observado nem sempre reflete com precisão as atitudes executadas externamente, já que os alunos podem ocultar suas emoções

ou respostas motivadas por fatores externos. Além disso, a observação por si só não é capaz de capturar completamente o contexto emocional ou psicológico do aluno, dificultando a análise de atitudes complexas. Essas limitações tornam o processo de avaliar atitudes mais subjetivo e excessivamente dependente da interpretação do avaliador, em nossa opinião. Logo concluímos que a avaliação das atitudes dos alunos requer a utilização de ferramentas mais específicas.

4.6 ENTREVISTAS COM OS PRINCIPAIS AUTORES INTERNACIONAIS

Nesta subseção trazemos as análises das entrevistas dos 4 autores/pesquisadores em Avaliação na área de Ensino de Química. Apresentamos as respostas dos quatro participantes para cada pergunta e discutimos temas semelhantes e divergentes que emergiram quanto às suas percepções.

Ao questionarmos os pesquisadores: “O que você entende por “avaliação”?”, estes declararam:

Acho que tenho um entendimento bem tradicional, e que avaliação é o processo de coletar informações do contexto que me interessa dos alunos. [...] Acho que no campo, na prática, as pessoas entendem a avaliação como isso. E quando as pessoas dizem avaliação, acho que muitas vezes querem dizer “evaluation”¹⁰¹ e não “assessment”¹⁰². Certo. Onde eles estão pegando alguns dados e submetendo-os a algum tipo de critério. Como talvez seja uma chave de resposta ou uma rubrica ou algo assim. Mas eu gosto de olhar para essas coisas separadamente. E então quando penso em avaliação, penso na ação de **coletar dados para tirar alguma conclusão sobre [a aprendizagem] do aluno** (P1, grifo nosso, tradução nossa).

Eu diria **coletar evidências** e fazer inferências a partir delas sobre a qualidade de algo. Então, você sabe, há **muitos tipos diferentes de avaliação**. Você sabe, poderíamos avaliar o aprendizado do aluno, podemos avaliar programas, podemos avaliar pesquisas, podemos projetar avaliações. Bem, de qualquer forma, acho que entraremos nisso, mas acho que como uma ideia abrangente, sim. Usar

¹⁰¹ “Evaluation” se concentra em fazer julgamentos sobre a qualidade ou valor de algo, como um programa, curso ou o desempenho do aluno. É tipicamente somativa e geralmente ocorre no final de uma unidade, curso ou ano acadêmico. Os métodos de avaliação podem incluir testes padronizados, exames finais e avaliações de desempenho. O propósito é determinar se os objetivos foram atingidos e informar decisões sobre notas, eficácia do programa ou melhorias futuras.

¹⁰² “Assessment” foca principalmente em reunir informações sobre o aprendizado e o progresso dos alunos e também é considerado uma linha/campo de pesquisa. Ela pode ser formativa e contínua, ajudando os professores a entender o quão bem os alunos estão aprendendo o material, por exemplo. Os métodos de avaliação podem incluir questionários, observações, discussões e projetos. O objetivo é fornecer feedback para melhorar o aprendizado e as estratégias instrucionais.

evidências para **fazer inferências** sobre a qualidade do que quer que você esteja avaliando (P2, grifo nosso, tradução nossa).

Avaliação é o processo de medir a compreensão do aluno e inclui os instrumentos que usamos para medir a compreensão do aluno (P3, tradução nossa)

Uma ampla gama de habilidades STEM¹⁰³ (Ciência, Tecnologia, Engenharia e Matemática) é essencial para que os alunos prosperem nas indústrias, ambientes e sociedade de hoje. Essas habilidades abrangem investigação científica, análise, trabalho em equipe, argumentação, pensamento crítico, resolução de problemas, inovação, criatividade e alfabetização em ciência e tecnologia. Pesquisas indicam que métodos de ensino e avaliação estão intimamente ligados ao desenvolvimento dessas habilidades STEM. Espera-se que os educadores integrem habilidades do século XXI em suas práticas de ensino e as passem aos alunos. A avaliação eficaz **vai além dos testes tradicionais**, promovendo aprendizagem significativa e habilidades de pensamento de ordem superior. (P4, grifo nosso, tradução nossa).

A partir da análise das respostas destacadas acima, observamos que há três semelhanças entre as percepções dos pesquisadores quanto às suas definições de avaliação. Essas semelhanças são: avaliações para tomada de dados; associação entre a avaliação e a realização de inferências/conclusões; e a pluralidade de métodos/tipos de avaliação.

Dois dos pesquisadores enfatizam a importância de coletar informações ou evidências dos alunos, destacando essa prática como um dos papéis centrais da avaliação. O P1, por exemplo, menciona “coletar dados”, e a P2 fala sobre “coletar evidências”. A respeito da associação entre a avaliação e a realização de inferências. A definição do P1 relata a importância de “tirar alguma conclusão sobre a [aprendizagem] do aluno”, enquanto o P2 se refere a “fazer inferências” sobre a qualidade de algo. Finalmente a respeito da pluralidade de tipos de avaliações, o P2 observa a diversidade de avaliações (aprendizagem dos alunos, programas, pesquisa), enquanto o P4 sugere a evolução dos métodos de avaliação além dos testes tradicionais, indicando uma conscientização sobre múltiplos tipos de avaliação.

A partir do exposto, notamos que os pesquisadores compartilham algumas ideias fundamentais sobre a avaliação, como as três supracitadas.

¹⁰³ STEM significa Ciência, Tecnologia, Engenharia e Matemática. Ele abrange uma ampla gama de campos e disciplinas que se concentram nessas áreas, promovendo o pensamento crítico, a resolução de problemas e a inovação. A educação STEM enfatiza o aprendizado prático e as aplicações do mundo real, visando preparar os alunos para carreiras nesses setores vitais, que são essenciais para o crescimento econômico e o avanço tecnológico (Pugliese, 2020).

Entretanto, notamos que suas percepções se diferenciam em outros pontos. A percepção do P4, por exemplo, nos parece mais abrangente, pois incorpora necessidades e habilidades educacionais contemporâneas, como: a investigação científica, a análise, o trabalho em equipe, a argumentação, o pensamento crítico, a resolução de problemas, a inovação, a criatividade e a alfabetização em ciência e tecnologia. Isso se distingue, principalmente das percepções do P1, que assume uma postura, caracterizado pelo próprio pesquisador como tradicional.

Ao questionarmos os pesquisadores: “Como você vê a avaliação como um campo de pesquisa no ensino de química?”, estes relataram:

Eu diria que estou feliz em ver mais pessoas examinando a avaliação. Acho que **temos um longo caminho a percorrer**. Acho que meio que arranhamos a superfície [estamos ainda na ponta do iceberg]. [...] E **é difícil para mim, eu acho, entender a avaliação sem currículo e instrução vinculados a isso**. E então eu acho que quando vejo a avaliação meio que retirada separadamente, onde não é **contextualizada** dentro do curso, dentro do currículo, dentro da instrução, acho que isso é problemático. Mas não estou apontando o dedo para outros pesquisadores, eu também sou culpada disso. Acho que é parte da nossa formação científica. Temos que escolher uma fatia muito, muito fina e então ir fundo. E então tentar pegar construções realmente complexas para provocar tudo em uma construção psicométrica separada e pensar que é isso que está acontecendo em espaços onde as pessoas estão ensinando e aprendendo ao mesmo tempo. Se olhássemos para toda a bagunça apenas em um grande globo, não seríamos capazes de fazer sentido algum nisso. Então eu acho que esse é o desafio de dar zoom in e zoom out, você sabe, meio que simultaneamente enquanto pensamos sobre problemas de pesquisa como sendo um dos grandes desafios que temos para a próxima geração de pesquisa em nosso campo... [...] então [se a avaliação] for muito estreita, então deixa de fora todas essas coisas importantes, mas se for muito ampla, então não somos capazes de dar zoom em qual é o conceito crucial. (P1, grifo nosso, tradução nossa)

Então, eu realmente não tinha pensado sobre isso como um campo de pesquisa, porque eu acho que **quase toda a educação em química está de alguma forma conectada à avaliação**. Se você está fazendo pesquisa, você sabe, você pode coletar todos os tipos diferentes de evidências para fazer inferências sobre isso. [...] Eu vejo quase tudo o que fazemos de alguma forma como avaliação [...], mas, se você me perguntasse o que estava mais intimamente relacionado à avaliação, eu suponho que seria isso, projetar e usar essas avaliações, esses instrumentos. Mas eu acho que nós poderíamos ampliar para incluir muitas outras coisas também. (P2, grifo nosso, tradução nossa).

A avaliação é **um foco crescente** no campo da educação em química. Esse crescimento é estimulado por mais pesquisadores formados em

teoria de medição e por avanços na tecnologia que estão ampliando o que é possível com a avaliação. (P3, grifo nosso, tradução nossa).

Este é um campo crucial na educação em química. Embora haja um corpo de pesquisa sobre avaliação de sala de aula de química e aprendizado de química, **é necessária uma investigação mais aprofundada** sobre a alfabetização e o conhecimento de avaliação dos professores de química. (P4, grifo nosso, tradução nossa).

A partir do exposto notamos que alguns pesquisadores indicam que ainda há muito a explorar e aprofundar na compreensão do que significa avaliar na área de Ensino de Química. O P1 menciona que ainda “temos um longo caminho a percorrer”. Similarmente o P3 enfatiza que a avaliação é um “foco crescente” no Ensino de Química. O P4 menciona que “é necessária uma investigação mais aprofundada” da avaliação em temas específicos como alfabetização (científica) e o próprio conhecimento de avaliação dos professores.

Além disso, notamos que o P1 e o P2 veem a avaliação como algo que está inserido em um contexto educacional maior/mais amplo, possuindo relações estreitas com outras dimensões. Tanto o P1 como o P2 destacam a inter-relação entre avaliação, currículo e instrução. O P1 discute os desafios de contextualizar a avaliação no contexto educacional e de pesquisa, queixando-se de um excesso de fragmentação. Para o P1, é difícil visualizar a avaliação sem estar “contextualizada” em um curso, currículo ou instrução. Nesse sentido, o P2 sugere que grande parte do Ensino de Química envolve inerentemente a avaliação ao afirmar que “quase toda a educação em química está de alguma forma conectada à avaliação”.

Apesar dessas semelhanças, notamos que as percepções dos pesquisadores a respeito da Avaliação na área de Ensino de Química como campo de pesquisa apresentam diferenças notáveis. O P1 enfatiza a complexidade da avaliação e a necessidade de estar contextualizada. O P2 compreende a avaliação como parte essencial e integrante do Ensino de Química. O P3 indica a importância dos avanços tecnológicos como catalisadores das transformações vistas. O P4 reconhece o papel crucial da avaliação como campo de pesquisa, porém advoga em favor de investigações que expandam nosso conhecimento de habilidades e as compreensões dos professores sobre avaliação.

Ao questionarmos os pesquisadores: “O que tem sido frequentemente avaliado, talvez excessivamente, no Ensino de Química. Por que você acredita que esse é o caso?”, estes relataram:

Então, ideias cientificamente adequadas dos alunos versus suas **ideias inadequadas**. Acho que uma vez que o modelo de mudança conceitual atingiu a educação científica no final dos anos oitenta. Especialmente porque as pessoas foram capazes de descobrir como fazer medições e ferramentas de medição bastante decentes para isso. [...]. Acho que construir inventários conceituais é um trabalho interessante, mas não houve aceitação suficiente do uso desses inventários e os efeitos de coletar esse tipo de dados de avaliação e ver o que você pode fazer com esses dados, curricularmente ou instrucionalmente, para avançar mais nesse campo. Acho que não precisamos de mais desse tipo de trabalho. [...] acho que continuaremos vendo mais e mais iterações desse mesmo tipo de coisa, que são inventários de conceitos, medindo ideias adequadas e inadequadas dos alunos para então ser capaz de fazer a alegação de que eles [os alunos] entendem isso, ou eles não entendem isso. Acho que, não conseguimos ir além disso neste momento, mas deveríamos. (P1, grifo nosso, inserção nossa, tradução nossa).

Ah, **concepções alternativas ou inadequadas**. O que os alunos não sabem. Acho que é absolutamente excessivo. Há milhares e milhares e milhares de artigos sobre isso em educação em química e educação em ciências. [...] está desaparecendo um pouco, mas ainda vejo muitos artigos surgindo, analisando o que os alunos não sabem. Não acho isso produtivo de forma alguma. Já fizemos isso até a morte. **O que é muito mais importante, eu acho, é descobrir o que os alunos sabem, não o que eles não sabem**. E acho que é o caso porque é fácil, na verdade é muito fácil fazer os alunos cometerem erros. (P2, grifo nosso, tradução nossa).

Eu não diria que é excessivamente avaliado, mas uma abordagem frequente na literatura é usar exames de múltipla escolha para medir o conhecimento e as habilidades algorítmicas dos alunos. (P3, tradução nossa).

A avaliação em química é focada no aprendizado e no ensino principalmente do ponto de vista do aluno. (P4, tradução nossa).

A partir dos excertos expostos, observa-se certa semelhança entre as respostas dos pesquisadores é a ênfase excessiva da pesquisa em investigar compreensões/concepções inadequadas ou equívocos dos alunos. Tanto o P1 como o P2 expressam uma frustração com o foco excessivo na avaliação de ideias inadequadas dos alunos. Já o P3 assume uma postura mais neutra, relatando que, exames de múltipla escolha são comumente usados, porém não excessivamente. O P4, por outro lado, relata que as avaliações tem sido centradas majoritariamente na perspectiva dos alunos. Inferimos que o P4 talvez deseje que a avaliação também seja centrada em outros pontos do contexto educacional, como na atividade docente (instrução, planejamento, performance, práticas, etc.) ou no currículo.

Ao questionarmos os pesquisadores: “O que não tem sido avaliado o suficiente no Ensino de Química? Por que você acredita que esse é o caso?”, estes relataram:

Acho que temos a oportunidade de estudar para tentar entender melhor **o pensamento dos alunos**. Então, acho que em química, houve muita influência da comunidade de pesquisa em educação em física, onde os alunos constroem modelos imprecisos do mundo só de estarem vivendo nele. Mas acho que em química, eu diria que a maioria dos modelos que não funcionam cientificamente para eles tendem a vir da instrução ou dos professores ou apenas do ambiente de aprendizagem. E acho que quero entender melhor, como podemos ajudar os alunos a construir ideias mais precisas? E só saber o que, quando eles erram uma resposta, não é suficiente. Gostaria de saber mais sobre como eles estão raciocinando sobre os conceitos de química? Então, acho que **deveríamos coletar mais dados sobre o pensamento dos alunos**. (P1, grifo nosso, tradução nossa)

O impacto dos currículos nos resultados dos alunos. [...] Há muitos artigos mostrando que se os ensinamos da mesma maneira tradicional, eles acabam não sabendo muito. Então, acho que projetar e realizar estudos longitudinais sobre o impacto das intervenções instrucionais não tem sido [realizado]. E a razão, novamente, é que é difícil. Você precisa de um estudo de longo prazo, precisa de financiamento de longo prazo, precisa ser capaz de acompanhar os alunos. E normalmente não fazemos isso no ensino superior. (P2, tradução nossa)

Não há avaliação suficiente que reflita “**o que os químicos realmente fazem**”. Isso pode incluir a criatividade necessária para resolver novos problemas ou a seleção de informações ou procedimentos relevantes para resolver novos problemas. Por exemplo, um químico praticante pode precisar isolar e caracterizar poluentes em uma amostra de sujeira. Fazer isso requer mais do que apenas uma execução em um HPLC¹⁰⁴. Nossa avaliação reflete isso? De forma mais geral, descobrimos que os alunos podem demonstrar conhecimento de fatos de química e procedimentos algorítmicos, mas não são treinados para saber quando essas informações são aplicáveis. Como resultado, elas são aplicadas em situações em que não eram pretendidas. Um exemplo é treinar alunos para converter a massa de um composto em número de mols. Observamos alunos em laboratório calculando os mols de um soluto pegando a massa da solução e dividindo-a pela massa molecular do soluto. É perfeitamente lógico, a partir de como eles foram treinados e avaliados; precisamos de uma avaliação que dê aos alunos feedback sobre situações como essas. (P3, tradução nossa)

A pesquisa necessária está no campo da investigação da **motivação dos alunos** e da escolha de carreira, bem como do aspecto do professor, especialmente do **conhecimento de avaliação**. (P4, grifo

¹⁰⁴ Cromatografia Líquida de Alta Eficiência (HPLC) é uma técnica de cromatografia em coluna que bombeia uma amostra em um solvente (fase móvel) através de uma coluna com fase estacionária.

nosso, tradução nossa)

A partir do exposto, nota-se que lacunas na avaliação do Ensino de Química são identificadas por cada um dos pesquisadores. Observa-se que o P1 enfatiza a avaliação da compreensão do aluno, expressando uma preocupação por uma visão mais profunda. Evidências disso são vistas quando o P1 relata que precisamos entender melhor “o pensamento do aluno”. Já o P2 acredita que necessitamos de investigações que avaliam os “impactos dos currículos”. O P3, por outro lado destaca a importância de entendermos melhor a aplicação do conhecimento dos alunos em contextos do mundo real, que espelham “o que os químicos realmente fazem”. Finalmente, o P4 acredita ser mais importante a investigação/avaliação da “motivação dos alunos” e o “conhecimento de avaliação dos professores”. Logo cada pesquisador destaca diferentes aspectos que necessitam de mais pesquisas acerca dessa temática.

Os resultados de nossa tese identificam uma predominância de artigos que avaliam sujeitos e/ou contextos do Ensino Superior em Ensino de Química. Ao questionarmos os pesquisadores sobre a predominância desse nível de ensino nos estudos que investigam a avaliação no Ensino de Química e possíveis razões para explicar esse fenômeno, os pesquisadores relataram:

Acho que não tenho certeza, além de muitas pessoas que podem se importar com esse trabalho serem docentes. E então pode haver essa sobreposição entre o papel de ser um professor de química e o papel de ser um pesquisador de educação em química. E então **eles têm mais acesso a seus alunos** e os alunos de outros professores. E então eu acho que é uma razão pela qual eu acho que muitas pessoas que estão no meu tipo de posição [docente e pesquisadora], elas não estudam educação K12¹⁰⁵ nos EUA porque **elas não têm acesso a alunos K12**. Mas eu estou mais interessada em estudar muitos ambientes K12. E então eu tenho que construir essas pontes. Mas esse é apenas o meu interesse. (P1, grifo nosso, inserção nossa, tradução nossa)

[...] eu vivo em um departamento de química e nós tendemos a estudar as coisas em que temos mais experiência. [...] Eu realmente **não tenho acesso a alunos pré-faculdade**. Então os pesquisadores de química têm acesso a alunos de ensino superior e eles tendem a se concentrar nisso. (P2, grifo nosso, tradução nossa).

Presumivelmente, isso significa que houve menos atenção dada à avaliação de química no ensino médio. No sistema dos EUA, acredito

¹⁰⁵ K-12 refere-se ao sistema educacional nos Estados Unidos que inclui todos os níveis de educação, do jardim de infância até a 12ª série (fim do Ensino Médio).

que tenha a ver com a forma como a pesquisa em educação em química está situada. A pesquisa em educação em química geralmente está localizada no departamento de química ou ciências naturais e eles são encarregados de avançar o que é conhecido no nível de ensino superior. **As colaborações necessárias para coletar dados no nível secundário não são desenvolvidas** nesses departamentos na medida em que são nos departamentos de educação. Nos departamentos de educação, há pesquisadores em educação científica que estudam avaliação em aulas de ciências no nível secundário, mas podem não ter foco em química. (P3, grifo nosso, tradução nossa).

Nas últimas duas décadas, há dois motivos para essa tendência:

A. Conduzir pesquisas em salas de aula se tornou muito mais complexo devido à necessidade de aprovações éticas e consentimento dos pais. Portanto, **é mais fácil obter aprovações de pesquisa ao estudar alunos com mais de 18 anos.**

B. Além disso, o campo da educação química inclui mais pesquisadores cuja pesquisa inicial foi em química, e somente mais tarde em suas carreiras eles fizeram a transição para a educação em química. Eles têm mais experiência em ensinar alunos universitários do que alunos do ensino médio. (P4, grifo nosso, tradução nossa).

A partir do exposto, notamos que todos os pesquisadores relataram que uma das principais razões para os artigos em Avaliação na área de Ensino de Química terem sido realizados, majoritariamente, no Ensino Superior é a facilidade de acesso a esse contexto. Os pesquisadores mencionam os desafios de acessar alunos da Educação Básica. O P1 enfatiza a dificuldade em estudar alunos do nível K12 devido à falta de acesso direto, enquanto P2 e P3 observam que pesquisadores no ensino superior normalmente não possuem acesso fácil a ambientes pré-universitários.

Além disso nota-se que alguns pesquisadores, principalmente o P3 e o P4, associam esse fenômeno na literatura com o contexto institucional, cujo os autores pertencem. O P3 e o P4 apontam que muitos dos pesquisadores de educação em química estão situados em departamentos de química ou ciências naturais, o que naturalmente leva a um foco em ambientes de ensino superior.

Ao questionarmos os pesquisadores sobre o interesse crescente por pesquisas em Avaliação na área de Ensino de Química nas últimas duas décadas, estes relataram:

Sim, acho que **houve muitas oportunidades oferecidas à pesquisa em educação em química**, bem como à pesquisa em educação em biologia **feita pela comunidade de pesquisa em educação em física**. [...] Acho que começamos a perceber que os alunos podiam

fazer coisas em um teste e ainda não saber muito, e acho que a pesquisa em educação em física foi pioneira nisso. E então, começamos a nos perguntar, ok, não é apenas sobre o que os alunos sabem. **Precisamos ter boas maneiras de medir o que eles sabem** para que nossas alegações sobre o que eles sabem sejam mais válidas. E então houve alguns trabalhos iniciais na pesquisa em educação em química. Foi **Mary Nakhleh** que fez um artigo sobre resolução de problemas versus compreensão de conceitos. [...] E então houve também alguns trabalhos de [Susan] **Nurrenbern** e [Miles] **Pickering**. Novamente, isso foi nos meados do final dos anos 1980. E esse trabalho revelou que os alunos podiam obter notas muito boas em testes, mas então, ao fazer tarefas mais conceituais ou em uma entrevista onde eles tinham que explicar seu raciocínio, eles não entendiam química. [...] Eu diria que esses trabalhos talvez tenham feito esse gráfico mudar. [...] Então eu me pergunto se há também algo sobre a **proliferação de pesquisadores** de educação em química dentro dos departamentos de química, se isso tem algum tipo de impacto nessas coisas também, porque assim que tivemos mais professores que estavam orientando alunos de doutorado e contratando pós-doutorados para realmente fazer esse trabalho, na disciplina de química, eu imagino que isso também teria sido um impulsionador. (P1, inserções nossas, grifo nosso, tradução nossa)

Bem, **há mais pessoas fazendo** isso. Acho que isso é parte. Quando eu comecei, havia poucas pessoas fazendo pesquisa em educação química. [...] E isso é porque houve um crescimento real de pessoas, eu acho, empregadas em departamentos de química como professores fazendo pesquisa em educação química nos últimos 20 anos, e isso se acelerou. [...] mas acho que se você olhar para o número de artigos em pesquisa em educação química, você verá que eles também cresceram em geral. [...], e há mais pessoas que precisam, que entendem a necessidade de medições confiáveis. [...] Então eu diria que provavelmente o maior impacto nos últimos 15 anos foi o **National Academy Report on Discipline Based Education Research**. Então, este é um relatório de consenso. Foi publicado em 2012, eu acho. Então, doze anos atrás, o que meio que deu o selo de aprovação à ideia de pesquisa educacional baseada em uma disciplina [...]. Havia mais programas de pós-graduação começando provavelmente há cerca de 20 anos do que antes, que formavam alunos de pós-graduação, que formavam pesquisadores que não eram apenas químicos, mas também tinham formação teórica em métodos educacionais. Então, eu acho que como um campo, a pesquisa educacional em química ainda é muito jovem e está amadurecendo. E conforme amadureceu, [...] se você vai medir algo, você tem que medir com **um instrumento que seja confiável** e não apenas algo que você inventou, você sabe, no verso de um envelope, que é o que costumava acontecer antes. [...] Muito mais departamentos desde a publicação daquele relatório fizeram um lugar para a educação em química de uma forma robusta. [...] No departamento de química, eles geralmente não eram professores, ou eram professores que não recebiam muito respeito, recursos ou financiamento, o que também é um grande negócio. E se você não tem alunos de pós-graduação, por exemplo, é difícil fazer pesquisa, sabe, e ter uma posição de docente em tempo integral, [...]. A outra coisa que fez a diferença é que a **National Science Foundation** fez um esforço particular para apoiar a pesquisa,

a pesquisa em educação em química no ensino superior. (P2, inserções nossas, grifo nosso, tradução nossa).

Acredito que essa tendência reflete o **crescimento geral da pesquisa em Educação em Química** nas últimas duas décadas. O número total de artigos da pesquisa em Educação em Química publicados provavelmente segue uma tendência semelhante. Essa tendência provavelmente é atribuída ao **número crescente de pesquisadores** de educação em química no campo e mais periódicos para compartilhar a pesquisa em Educação em Química. (P3, grifo nosso, tradução nossa).

A realização de pesquisas em avaliação se tornou mais frequente devido a **estudos internacionais como o PISA** e aos **líderes universitários** que estão pressionando por currículos, habilidades e certificados padronizados. (P4, grifo nosso, tradução nossa).

A partir do exposto, nota-se que três pesquisadores (P1, P2 e P3) relataram o aumento no número de indivíduos envolvidos em pesquisa educacional em química nas últimas duas décadas como um dos principais fatores para explicar o crescente interesse. De acordo com os entrevistados, esse crescimento de pesquisadores é visto como uma força motriz por trás do crescente interesse em avaliação.

Além disso, os pesquisadores P1 e P2 relatam que uma necessidade de medições/avaliações mais confiáveis pode ter contribuído para o aumento do número de artigos. Nesse, sentido os pesquisadores relataram que, à medida que a área evoluiu, houve uma maior consciência da necessidade de ferramentas de medição eficazes para avaliar o aprendizado do aluno com precisão.

Apesar dessas semelhanças, encontramos diferenças notáveis entre as percepções dos pesquisadores sobre o crescente interesse por pesquisas em Avaliação na área de Ensino de Química. Para o P1, isso pode ser atribuído as percepções obtidas da comunidade de pesquisa em Educação em Física, que influenciaram a área de Ensino de Química. Isso é evidente na fala: “houve muitas oportunidades oferecidas à pesquisa em educação em química, bem como à pesquisa em educação em Biologia feita pela comunidade de pesquisa em Educação em Física”

O P1 também mostra a percepção de que as avaliações tradicionais podem não refletir com precisão a compreensão dos alunos. Isso é evidenciado na fala: e “Precisamos ter boas maneiras de medir o que eles [os alunos] sabem”. Além disso o P1 aponta alguns autores como figuras influentes para o aumento do interesse

em pesquisas sobre avaliação, como por exemplo: Mary Nakhleh, Susan Nurrenbern e Miles Pickering.

Por outro lado, o P2 considera a aceitação da pesquisa educacional como uma área válida de investigação dentro da disciplina de Química como um fator relevante. Além disso, o P2 menciona o relatório da *National Academy Report on Discipline Based Education Research* como um fator importante. O P2 também discute que a Educação em Química é um campo relativamente jovem, que ainda está amadurecendo, o que implica no crescimento e sofisticação dos métodos de avaliação disponíveis. De acordo com o P2 isso pode estar relacionado ao interesse por pesquisar o tema em Ensino de Química. Apesar dos fatores supracitados, o P2 acredita que o crescimento no interesse não é exclusivo para o tema avaliação, apenas. O P2 acredita que esse fenômeno é visto de forma abrangente em todas as pesquisas em Ensino de Química.

Já o P3 e P4, optam por ser mais breves. O P3 enfatiza o aumento geral de artigos do Ensino de Química e relaciona isso com o estabelecimento de mais periódicos para compartilhar essas pesquisas, sugerindo que a infraestrutura para disseminar pesquisas cresceu. O P4 destaca pressões externas feitas por líderes universitários e estudos internacionais, como o PISA que impulsionaram a necessidade de avaliações padronizadas.

Assim, a maioria dos pesquisadores reconhece um interesse crescente em pesquisas em Avaliação na área de Ensino de Química, impulsionado pelo aumento na atividade de pesquisa oriundo do maior número de indivíduos trabalhando na área. Também relatam a necessidade de métodos de avaliação confiáveis como um fator influenciador. Os pesquisadores divergem ao oferecerem suas perspectivas sobre os fatores mais específicos que contribuíram para essa tendência. Dentre eles, são mencionadas influências históricas, suporte institucional, o contexto mais amplo da pesquisa educacional, e autores e relatórios específicos. Essas percepções demonstram a complexidade da área e os diversos aspectos que influenciam na formação das práticas de pesquisa atuais de Avaliação na área de Ensino de Química.

Ao questionarmos os pesquisadores sobre o que tem sido avaliado em Ensino de Química, por meio da apresentação do Quadro 18 com as dez categorias e suas distribuições, em porcentagens, no *corpus*, estes relataram:

Então, quando olho para este quadro, **não necessariamente acho que seja muito específica para química**. Em outras palavras, se olhássemos para isso para outra ciência natural, estou me perguntando se seria semelhante, tipo, para a geociência, geologia, física, biologia, biologia molecular, bioquímica ou química. [...] Acho que para várias dessas coisas, provavelmente veríamos essas mesmas coisas em disciplinas que não são nada parecidas com as ciências naturais. [...] Mas, novamente, não temos nada para comparar porque você não tem colegas em todas as disciplinas que fizeram a mesma revisão de literatura que você fez. (P1, grifo nosso, tradução nossa).

Acho que a categoria quatro... Então, essa que incluiria o design do instrumento e a avaliação dele... [...] Acho que é alta por causa do reconhecimento na comunidade de que, **coisas como essas são necessárias**. E, novamente, **são relativamente fáceis de fazer**. Na verdade, não é tão difícil fazer um novo instrumento. Se você usar as técnicas e métodos certos, pode fazer um instrumento para medir algo e pode fornecer alguma validade para ele. Se isso é significativo ou não é outra questão. [...] E há muitos argumentos sobre se isso também é válido, mas não importa. [...] [ao olhar para as outras categorias] Sim, bem, **o que vejo aqui é que há muita avaliação das coisas e menos avaliação dos resultados**... [...] menos impacto, menos observação, sabe, qual é o resultado real de todas essas coisas? [...] Como um campo, sabe, temos muitos instrumentos agora, há dezenas deles, centenas provavelmente, que você poderia escolher. Sabe, **não precisamos de mais instrumentos**. [...] **Acho que precisamos nos concentrar mais nos impactos reais**. [...] Nosso objetivo é ajudar os alunos a aprender e ajudá-los a se tornarem autossuficientes [...]. Eu não vejo muito disso acontecendo aqui. [...] Bem, acho que se você analisar [os artigos], verá que quase nenhum deles é de **longo prazo**. Quase todos eles estarão olhando para alguma coisa pequena. [...] Seu projeto é sobre o que o campo está valorizando na avaliação, mas também tenho um olhar muito cético sobre alguns dos artigos publicados que parecem passar pela revisão por pares. (P2, inserção nossa, grifo nosso, tradução nossa).

Acho que este quadro mostra as muitas maneiras pelas quais a avaliação é importante no campo e fala sobre a **crescente atenção ao rigor na medição** no campo. Por exemplo, **não é mais aceitável na literatura apontar anedotas** como evidência de eficácia curricular ou melhoria de atitudes dos alunos. Há uma expectativa de que estas devem ser sistematicamente avaliadas para serem confiáveis e isso se reflete nesta tabela. (P3, grifo nosso, tradução nossa)

Não tenho certeza se usaria o termo "propostas de ensino" como você fez. Eu diferenciaria entre avaliação da **perspectiva dos alunos** (desempenho, tarefas e assim por diante) e da **perspectiva dos professores** (conhecimento de avaliação dos professores). (P4, grifo nosso, tradução nossa).

A partir do exposto, nota-se que os pesquisadores possuem diferentes percepções a respeito das categorias. Para o P1, as categorias não

parecem ser específicas para a química. Nesse sentido, se um estudo semelhante fosse feito para outras disciplinas, relativas às Ciências Naturais, categorias semelhantes seriam geradas. Infelizmente, como o próprio P1 relatou, não temos parâmetros de comparação, pois “não tem colegas em todas as disciplinas que fizeram a mesma revisão de literatura que você fez”.

Ao analisarmos a resposta do P2, observamos que o pesquisador acredita que existe uma grande facilidade de desenvolver novos instrumentos de avaliação, o que explica a alta incidência de artigos alocados na categoria C04: Avalia um instrumento, ferramenta ou tecnologia. Nesse sentido, o P2 relata que a comunidade acadêmica em Ensino de Química já possui um grande número de ferramentas/instrumentos de avaliação, que talvez até seja excessivo. De acordo com o P2, a atenção da comunidade acadêmica em Ensino de Química deveria ser voltada para a avaliação das consequências e impactos no sentido mais amplo, possivelmente em estudos de longo prazo. Nesse sentido, acreditamos que o P2 possui um certo ceticismo em relação à validade de alguns dos estudos publicados, questionando a qualidade de algumas pesquisas desenvolvidas em Avaliação na área de Ensino de Química.

Já o P3 se expressa de forma mais breve, atribuindo os resultados (as categorias) a crescente importância do rigor na avaliação. De acordo com o P3 não se pode mais depender de evidências anedóticas para justificar/explicar as eficácias curriculares ou melhorias relatadas nas pesquisas. Logo, há uma necessidade crescente por avaliações sistemáticas, o que explica a necessidade de instrumentos adequados para medir esses aspectos. Assim, reconhecemos que o P3 se concentra, mesmo que não seja explicitamente relatado, na categoria C04: Avalia um instrumento, ferramenta ou tecnologia. Nesse sentido, semelhantemente ao P2.

O P4 opta por se concentrar na diferenciação entre a avaliação sob a ótica dos alunos e dos professores. Essa distinção sugere uma análise mais centrada nos sujeitos, no sentido de como eles percebem e utilizam as avaliações. Nesse ponto o P4 se diferencia dos outros pesquisadores por não considerar o que foi avaliado, mas quem foi avaliado. A respeito desse ponto, levantado pelo P4, indicamos um de nossos resultados. Uma das tendências identificadas após nosso movimento de revisão, mostrou a predominância de artigos que realizam a avaliação de alunos em comparação à avaliação de professores. De fato, observamos uma clara prioridade na avaliação dos alunos, com 68% dos artigos da categoria C05, que avalia

compreensões e percepções. Por outro lado, apenas 32% se concentraram nos professores.

Ao questionarmos os pesquisadores sobre os principais obstáculos e dificuldades da avaliação em sala de aula hoje, estes relataram:

Acho que **uma barreira para uma avaliação realmente reside na cabeça do instrutor**. Acho que se o instrutor tem ideias imprecisas de ensino e aprendizagem, ou seja, se ele está usando algum modelo de aprendizagem que não é compatível com a forma como os humanos realmente aprendem, então acho que ele não será capaz de avaliar de forma apropriada, autêntica e útil para o aluno. E então acho que isso remonta a um ponto anterior sobre olhar para a avaliação separadamente do ecossistema do ensino e aprendizagem [...] muitas dessas ideias precisam voltar e ser fundamentadas, por exemplo, como funciona a aprendizagem e o que conta como aprendizagem? E então podemos começar a dizer, ok, então como é a aprendizagem para esta unidade instrucional ou este curso? Certo. Como é a aprendizagem e quais podem ser os mecanismos sobre essa aprendizagem sobre os quais estamos fazendo suposições? E então como tornaríamos isso visível? E então quais tipos de dados podemos coletar para capturar isso? Sabe, eu sinto que tudo volta para algumas das nossas **suposições que estamos fazendo sobre o aprendizado**. E se o instrutor está fazendo suposições imprecisas, então a avaliação sempre será. [...] Então eu acho que isso sempre será uma barreira. Temos que focar se o **ecossistema de ensino e aprendizagem** não for construído sobre como os humanos aprendem, então eu não sei como realmente nos avançaremos. Então isso é uma barreira. (P1, grifo nosso, tradução nossa)

Vou lhe dizer como é aqui, e para muitas pessoas nos Estados Unidos agora está ficando **cada vez mais difícil solicitar dados**, dados de histórico sobre os alunos. [...] E então há uma governança cada vez mais forte concedida aos dados dos alunos, e temos que ser extremamente cuidadosos. Sei que parece que quero voltar aos bons e velhos tempos, mas, você sabe, temos conselhos de revisão institucional que governarão o que fazemos, e as coisas que costumávamos fazer, que era pedir permissão aos alunos para usar seus dados, essas não são mais boas o suficiente. [...] Acho que muitas das pesquisas que vemos são feitas por pesquisadores que ainda estão investigando aulas muito tradicionais. E então o que vemos é uma litania¹⁰⁶ do que os alunos não conseguem/são capazes de fazer. E os pesquisadores provavelmente **não têm a agência** para implementar um novo currículo e testá-lo completamente, por exemplo, na verdade, há poucos exemplos disso. [...] **Eu quero saber se posso ter um impacto** em como os alunos aprendem e entendem química. E apenas a mecânica de implementar mudanças nas salas de aula está ficando cada vez mais difícil, eu acho. Então você verá muitas pesquisas feitas nas salas de aula de outras pessoas em

¹⁰⁶ Uma litania é uma forma de oração ou invocação em que se repetem frases ou súplicas, geralmente em um formato responsivo. É comumente utilizada em contextos religiosos, mas também pode ser aplicada em rituais ou celebrações (Houaiss).

circunstâncias menos que ideais, digamos assim. (P2, grifo nosso, tradução nossa).

A avaliação em sala de aula precisa **medir características significativas** e precisa fazer isso com eficiência para dar **feedback oportuno a um grande número de alunos**. A tecnologia oferece algum potencial para atingir esse objetivo, mas geralmente é superestimada e requer desenvolvimento cuidadoso. Está claro que, daqui para frente, as próximas gerações de alunos terão acesso a interfaces de **IA** e isso mudará a forma como a resolução de problemas e a descoberta de conhecimento são conduzidas. O currículo precisa mudar para preparar os alunos para o uso de ferramentas baseadas em IA e a avaliação precisa mudar para fornecer feedback dos alunos sobre suas proficiências com essas ferramentas. (P3, grifo nosso, tradução nossa).

Quase **não há formação para futuros professores** e não há desenvolvimento profissional na área de **conhecimento de avaliação** de professores. (P4, grifo nosso, tradução nossa).

A partir do exposto, notamos que os pesquisadores consideram diferentes aspectos como os mais desafiadores. Contudo, uma aproximação das percepções do P1 e P4 pode ser feita. O P1 e o P4 abordam a falta de compreensão/conhecimento dos professores sobre como a aprendizagem realmente ocorre. O P1 discute que suposições inadequadas a respeito da aprendizagem, por parte dos professores pode afetar, negativamente, a eficácia das avaliações. Já o P4 relata a falta de formação específica dos professores, especialmente, relacionadas à avaliação.

O P2 considera a dificuldade crescente em obter dados sobre os alunos, como um dos principais desafios. De acordo com o P2, essa limitação institucional de governança impede que os pesquisadores compreendam o impacto de suas estratégias e inovações em diferentes grupos de alunos e contextos. Por esse motivo, o P2 relata que as pesquisas recentes tem se restringido a contextos tradicionais.

O P3 relata o desafio de incorporar ferramentas tecnológicas, como a inteligência artificial, tanto nas avaliações quanto no currículo. De acordo com o P3, a evolução das ferramentas que os alunos possuem acesso requer um ajuste nas avaliações, para que estas avaliem competências dos alunos no uso dessas novas tecnologias. O P3 também discute o desafio de fornecer feedback rápido e objetivo para um grande número de alunos. De fato, esse é um desafio encontrado pelo professor, principalmente devido ao número elevado de alunos em algumas turmas e

características da própria química, como a presença de conceitos abstratos, a aplicação da matemática e dificuldade de visualização/representação mental de moléculas, reações e fenômenos. O feedback precisa ser eficiente para que os alunos possam identificar e corrigir seus erros, e realizados de uma maneira para não sobrecarregar o professor. Nesse sentido, o feedback formativo, que ocorre durante o processo de aprendizagem e não aponta meramente os erros, porém também oferece sugestões de como melhorar, pode ser considerado uma estratégia válida, pois permite que o aluno ajuste seu desempenho ao longo do processo.

Em conclusão, as entrevistas com os quatro pesquisadores em Avaliação no Ensino de Química mostram a complexidade do tema e a diversidade de perspectivas. Apesar dos pesquisadores concordarem em alguns aspectos fundamentais a respeito da avaliação, como seu papel na coleta de dados e a promoção da utilização de diferentes métodos de avaliação, consideramos que as perspectivas diferentes, específicas a cada pesquisador refletem as particularidades e os desafios que tiveram que enfrentar, além dos contextos aos quais cada um pertence.

5 CONSIDERAÇÕES FINAIS

Dados os resultados desta investigação, retomamos nossas questões de pesquisa: I) Quais são as principais características, tendências e lacunas dos artigos internacionais sobre Avaliação na área de Ensino de Química? II) O que tem sido avaliado nos artigos internacionais sobre Avaliação na área de Ensino de Química? III) Quais são as perspectivas dos principais autores internacionais em Avaliação na área de Ensino de Química sobre as características, tendências e as lacunas das pesquisas?

A respeito das características, tendências e lacunas, notou-se um aumento significativo de artigos ao longo do recorte temporal deste estudo. Acreditamos que o aumento das publicações sobre avaliação na área de Ensino de Química, principalmente entre 2011 e 2021, indica uma evolução nas prioridades da comunidade acadêmica. Durante a primeira década analisada (2002-2011), apenas 18% dos artigos Avaliação na área de Ensino de Química foram publicados. Já na segunda década (2012-2021), esse número aumentou para 82%. Esse aumento significativo sugere uma maior valorização da avaliação como área de pesquisa em Ensino de Química.

Também atribuímos a introdução e a popularização do conceito de avaliação formativa, especialmente a partir dos estudos de Black e Wiliam para o aumento do interesse nesse campo. As avaliações formativas se diferenciam das avaliações tradicionais, por monitorarem o progresso dos alunos ao longo dos processos de ensino e aprendizagem, permitindo que os professores realizem ajustes pontuais na instrução, práticas e propostas de ensino. Assim, a abordagem também fomenta um ambiente de aprendizagem mais dinâmico e centrado no aluno, o que se alinha ao paradigma de ensino que valoriza a participação ativa dos estudantes em seu próprio processo de aprendizado.

A significativa predominância de publicações no *Chemistry Education Research and Practice* e *Journal of Chemical Education*, que representam 66% e 16% das publicações, respectivamente, destaca a importância dessas plataformas na disseminação de conhecimento e nas discussões sobre práticas avaliativas na comunidade de Ensino de Química. Esses periódicos se caracterizaram como essenciais para a divulgação de pesquisas sobre avaliação.

Outra característica relevante é a concentração de autores do estado da Flórida, nos Estados Unidos. A respeito de instituições específicas, a *University of South Florida* e a *Miami University* se destacam pela alta produtividade em publicações sobre avaliação. Isso se deve, em grande parte, à presença de pesquisadores influentes nessas instituições, como Ellen J. Yeziarski, Scott E. Lewis e Jennifer E. Lewis, que têm contribuído de maneira significativa para a literatura na área.

Quanto ao nível de ensino abordado, a maioria dos artigos focou no ensino superior, que representou 50% do total. Já a Educação Básica consistiu em apenas 8% do *corpus*. Além disso, muitos artigos também investigaram a avaliação de professores, refletindo uma visão mais abrangente sobre o papel da avaliação no contexto educacional. Essa atenção à avaliação docente é particularmente interessante, pois, embora a discussão sobre avaliação frequentemente se concentre na aprendizagem dos alunos, o reconhecimento da importância do professor no processo de ensino é fundamental.

A respeito do que foi avaliado nos artigos sobre Avaliação na área de Ensino de Química, resultou em dez categorias. Essas categorias foram: C01: Avalia habilidades (8,8%); C02: Avalia estratégias de avaliação, provas, exames e itens (8,3%); C03: Avalia propostas de ensino (19%); C04: Avalia um instrumento, ferramenta ou tecnologia (17,5%); C05: Avalia compreensões, percepções e interesse (9,3%); C06: Avalia conhecimentos (9,8%); C07: Avalia o currículo (6,8%); C08: Avalia atitudes (4,9%); C09: Avalia outras pesquisas (1,5%); e C10: Discute avaliação em contextos diferentes (14,1%).

A categoria mais proeminente foi a C03: Avalia propostas de ensino, que representou 19% do total de artigos, revelando a primeira tendência de pesquisa entre os artigos sobre Avaliação na área de Ensino de Química. A predominância dessa categoria demonstra um forte interesse dos pesquisadores em investigar a eficácia, a qualidade e o impacto de diversas abordagens pedagógicas, intervenções e atividades educativas. As propostas avaliadas abrangeram uma ampla gama de propostas, incluindo frameworks, workshops e sequências didáticas. Isso sugere que os pesquisadores não estão apenas implementando novas metodologias, mas também estão interessados em avaliar como essas abordagens influenciam o aprendizado dos alunos. Reconhecemos a necessidade de futuras pesquisas que não

apenas avaliem propostas individualmente, mas que também investiguem sua aplicabilidade em diferentes cenários educacionais.

A segunda tendência identificada refere-se à avaliação de instrumentos, ferramentas ou tecnologias, representada pela categoria C04: Avalia um instrumento, ferramenta ou tecnologia, que correspondeu a 17,5% do *corpus*. Isso revela uma crescente preocupação nas pesquisas em avaliar recursos tecnológicos e instrumentos/ferramentas avaliativas. Acreditamos que esse resultado está ligado à crescente conscientização a respeito da necessidade dos professores em integrarem ferramentas contemporâneas em suas práticas pedagógicas, especialmente devido ao acesso, cada vez mais precoce, dos alunos à tecnologia.

A terceira tendência observada foi a alta incidência de artigos na categoria C10: Discute avaliação em contextos diferentes, que representou 14,1% do total. Embora esses artigos não realizaram avaliações empíricas diretas de sujeitos ou objetos no contexto educacional, as discussões teóricas que desenvolveram foram essenciais para contextualizar a prática avaliativa, possibilitar a reflexão e o diálogo na comunidade acadêmica. Essa tendência sugere que, além de aplicar e avaliar, os pesquisadores buscam reflexões aprofundadas sobre as práticas de avaliação.

A quarta tendência identificada diz respeito à presença de artigos que realizam a avaliação tanto de alunos como de professores. Notou-se que 68% dos artigos da categoria C05, que avalia compreensões e percepções, avaliou alunos e 32% se concentraram nos professores. Reconhecemos a importância da avaliação de ambos os sujeitos envolvidos nos processos de ensino aprendizagem para compreender tanto como os alunos constroem o conhecimento, e como os professores desenvolvem suas práticas avaliativas. Isso aponta para uma preocupação em identificar as lacunas no entendimento dos alunos sobre os conteúdos, assim como os desafios enfrentados pelos professores ao aplicar propostas de ensino e avaliações. Essa característica do *corpus* indica uma análise mais completa das necessidades de formação dos professores e das dificuldades específicas de aprendizagem dos alunos.

Em relação às perspectivas dos principais autores internacionais em Avaliação na área de Ensino de Química, semelhanças e divergências entre suas perspectivas foram encontradas. Os pesquisadores concordam em três aspectos fundamentais sobre avaliação: coleta de dados, inferências, e pluralidade de métodos. Os pesquisadores destacaram o papel da avaliação em coletar dados dos alunos;

reconheceram o potencial da avaliação em permitir a tomada de decisão ou a realização de inferências sobre o aprendizado; e advogaram a favor da diversificação dos tipos de avaliação, indo além de testes tradicionais.

A respeito das lacunas das pesquisas em Avaliação na área de Ensino de Química, os pesquisadores mencionaram: uma necessidade de maior conhecimento a respeito das compreensões dos alunos (foco em compreender o que sabem e não o que não sabem); uma necessidade de estudos de longo prazo e/ou estudos que investigam os impactos reais na aprendizagem dos alunos; a aplicação do conhecimento dos alunos em contextos reais; a motivação/interesse dos alunos; e a formação de professores. Os pesquisadores também ofereceram explicações para a alta incidência (50%) de artigos que avaliaram o Ensino Superior. Os pesquisadores mencionaram a facilidade de acesso a esse público de alunos, devido a estes estarem frequentemente nas mesmas universidades que os próprios pesquisadores trabalham. Os pesquisadores também mencionaram as crescentes dificuldades para acessar alunos da Educação Básica e utilizá-los como sujeitos de pesquisa.

A respeito do aumento de interesse por pesquisas em Avaliação na área de Ensino de Química, a maioria dos pesquisadores mencionou, direta ou indiretamente, o aumento no número de pesquisadores na área como o fator principal que impulsionou esse fenômeno. Especificamente nesse ponto, o P1 e o P2 também associaram o fenômeno à crescente necessidade de medições mais confiáveis, que impulsionou as pesquisas, enquanto o P3 e o P4 relacionaram o crescimento do interesse à melhor infraestrutura para disseminação de pesquisas. A respeito dos principais desafios encontrados em avaliar atualmente, os pesquisadores discutiram: a falta de compreensão dos professores sobre como a aprendizagem ocorre; as dificuldades em obter dados dos alunos; e a necessidade de adaptar avaliações para incorporar novas tecnologias.

Em conclusão, os resultados desta pesquisa demonstram a transformação significativa nas publicações sobre Avaliação na área de Ensino de Química entre 2002-2021. As características das publicações ao longo do recorte temporal sinalizam um crescente reconhecimento da relevância desse tema de pesquisa. O aumento expressivo de artigos, particularmente após 2011, aliado à adoção do conceito de avaliação formativa pela comunidade, aponta para uma mudança paradigmática que prioriza o acompanhamento dos processos de aprendizagem dos alunos e a adaptação das práticas pedagógicas. As dez categorias

elaboradas a partir dos resultados dessa pesquisa mostram um panorama do que tem sido valorizado pela comunidade acadêmica nos últimos anos. As perspectivas dos principais autores internacionais enfatizam as necessidades que consideram mais urgentes, e apontam para direções promissoras para futuras investigações. É relevante que futuras pesquisas em Avaliação na área de Ensino de Química abordem os pontos levantados pelos autores entrevistados e fundamentem seus estudos nas características, tendências, lacunas e categorias apresentados nessa pesquisa.

Embora esta pesquisa tenha proporcionado uma visão abrangente das características, tendências e lacunas da temática Avaliação na área de Ensino de Química, bem como as percepções de alguns dos principais autores, existem algumas limitações que devem ser destacadas. Uma delas foi a escolha das bases de dados utilizadas. Optamos por não incluir outras bases de dados relevantes, como o *Web of Science*, devido ao aumento significativo no volume de resultados, o que tornaria inviável a execução das etapas da revisão bibliográfica sistemática de uma amostra tão extensa dentro do tempo limitado disponível para a pesquisa. Reconhecemos que esse fator pode ter restringido o alcance da análise, deixando de fora artigos importantes que poderiam enriquecer ainda mais as análises. Avaliação na área de Ensino de Química. Para futuras investigações, recomendamos a inclusão desta base de dados e possivelmente outras, possibilitando uma análise mais exaustiva da literatura.

Outra limitação diz respeito ao Modelo do Triângulo da avaliação para as categorias do currículo e das atitudes realizado nos Resultados dessa tese. Embora o triângulo tenha sido uma ferramenta útil, percebemos que ele apresenta limitações para uma análise mais profunda de alguns aspectos específicos. A estrutura do triângulo, ao se concentrar na cognição, observação e interpretação, não é suficiente para capturar a complexidade da dinâmica curricular e de atitudes, representando nossas categorias C08 e C09, respectivamente. Essa limitação aponta para a necessidade de um aprofundamento da investigação sobre a adaptação, aprimoramento dessa estrutura, considerando diferentes enfoques teóricos e metodológicos. Seria interessante que futuros estudos explorassem esse modelo, expandindo-o.

Também, um movimento que ficou de fora neste estudo, mas que seria de grande valor, é a elaboração de uma cronologia das referências mais utilizadas pelos autores dos artigos analisados. Com o mapeamento das referências

utilizadas Avaliação na área de Ensino de Química seria possível identificar os momentos e as publicações decisivas para o desenvolvimento da área no contexto internacional. Este movimento também iria contribuir para o entendimento das principais influências que moldaram as práticas avaliativas na área, ajudando a construir uma linha do tempo do seu desenvolvimento.

Em conclusão, esta pesquisa contribui para o entendimento das principais características, tendências e lacunas sobre o tema Avaliação na área de Ensino de Química, fornecendo uma visão detalhada da evolução da literatura no período de 2002 a 2021. Mesmo diante das limitações já mencionadas, consideramos que os resultados obtidos são valiosos para compreender o que a área valoriza nas pesquisas sobre avaliação. Além disso, consideramos que as perspectivas dos principais autores internacionais refletem as características observadas e são indicativos claros de que a Avaliação na área de Ensino de Química continua em expansão e desenvolvimento.

As sugestões para pesquisas futuras, como a ampliação das bases de dados, a expansão do Modelo do Triângulo da avaliação e a criação de uma cronologia das referências utilizadas pelos autores dos artigos analisados, somam-se a outras possibilidades de investigações que poderão oferecer novas perspectivas e direcionamentos para os desafios encontrados na avaliação, de um modo geral, e mais especificamente no Ensino de Química.

REFERÊNCIAS

- ABRAHAM, M. R.; VISE, K. D.; PELLEGRINO, J. W. Teaching and learning science through inquiry: an overview. **Journal of Science Teacher Education**, v. 5, n. 2, p. 83-93, 1994.
- ACHIEVE. **Next generation science standards:** for states, by states. 2016. Disponível em: <https://www.nextgenscience.org>. Acesso em: 4 nov. 2024.
- ALI, N.; ULLAH, S.; RAEES, M. Interactive cube for effective demonstration of virtual periodic table. **Education and Information Technologies**, v. 27, n. 2, p. 1635-1654, 2022.
- ALLENBAUGH, R. J.; HERRERA, K. M. Pre-assessment and peer tutoring as measures to improve performance in gateway general chemistry classes. **Chemistry Education Research and Practice**, v. 15, n. 4, p. 620-627, 2014.
- ALMEIDA, P. A. et al. The interplay between students' perceptions of context and approaches to learning. **Research Papers in Education**, v. 26, n. 2, p. 149-169, 2011.
- ANTONOGLOU, L. D.; CHARISTOS, N. D.; SIGALAS, M. P. Design, development and implementation of a technology enhanced hybrid course on molecular symmetry: students' outcomes and attitudes. **Chemistry Education Research and Practice**, v. 12, n. 4, p. 454-468, 2011.
- ANWAR, Y. A. S.; MUTI'AH, M. A. Exploration of the scientific papers and self-assessment of students using the COVID-19 case on biochemistry course. **Biochemistry and Molecular Biology Education**, v. 49, n. 3, p. 326-332, 2021.
- APUGLIESE, A.; LEWIS, S. E. Impact of instructional decisions on the effectiveness of cooperative learning in chemistry through meta-analysis. **Chemistry Education Research and Practice**, v. 18, n. 1, p. 271-278, 2017.
- ARMSTRONG, L. B. et al. Developing a green chemistry focused general chemistry laboratory curriculum: what do students understand and value about green chemistry?. **Journal of Chemical Education**, v. 96, n. 11, p. 2410-2419, 2019.
- ATIYEH, E. L.; YORK, D. M.; MUÑIZ, M. N. Beneath the surface: an investigation of general chemistry students' study skills to predict course outcomes. **Journal of Chemical Education**, v. 98, n. 2, p. 281-292, 2020.
- ATKINS, P. W.; FRIEDMAN, R. **Molecular Quantum Mechanics**. Oxford: Oxford University Press, 2011.
- AYDIN, S. A science faculty's transformation of nature of science understanding into his teaching graduate level chemistry course. **Chemistry Education Research and Practice**, v. 16, n. 1, p. 133-142, 2015.
- AYDIN, S.; BOZ, Y. The nature of integration among PCK components: A case study of two experienced chemistry teachers. **Chemistry Education Research and Practice**, v. 14, n. 4, p. 615-624, 2013.

AYDIN, S.; FRIEDRICHSEN, P. M.; BOZ, Y.; HANUSCIN, D. L. Examination of the topic-specific nature of pedagogical content knowledge in teaching electrochemical cells and nuclear reactions. **Chemistry Education Research and Practice**, v. 15, n. 4, p. 658-674, 2014.

BAILIN, S. Critical thinking and science education. **Science Education**, v. 11, p. 361–375, 2002.

BALAKRISHNAN, S. Who Wants to Be a Chemist? Formative Assessment Made Fun and Engaging. **Journal of Educational Technology**, v. 15, n. 2, p. 1-7, 2018.

BARDIN, L. **Análise de conteúdo**. 3. ed. Lisboa: Edições 70, 2011.

BEHMKE, D. A.; ATWOOD, C. H. Implementation and assessment of Cognitive Load Theory (CLT) based questions in an electronic homework and testing system. **Chemistry Education Research and Practice**, v. 14, n. 3, p. 247-256, 2013.

BEKTAS, O. et al. Pre-service chemistry teachers' pedagogical content knowledge of the nature of science in the particle nature of matter. **Chemistry Education Research and Practice**, v. 14, n. 2, p. 201-213, 2013.

BENNETT, R. E. Formative assessment: a key to deeper learning. **Educational Leadership**, v. 69, n. 6, p. 28-33, 2011.

BERNHOLT, S.; PARCHMANN, I. Assessing the complexity of students' knowledge in chemistry. **Chemistry Education Research and Practice**, v. 12, n. 2, p. 167-173, 2011.

BIGGS, J.; TANG, C. Train-the-trainers: implementing outcomes-based teaching and learning in Malaysian higher education. **Malaysian Journal of Learning and Instruction**, v. 8, p. 1-19, 2011.

BLACK, P.; WILIAM, D. Assessment and classroom learning. **Assessment in Education: Principles, Policy & Practice**, v. 5, n. 1, p. 7-74, 1998.

BLACK, P.; WILIAM, D. Assessment for learning in the classroom. In: GARDNER, J. (Ed.). **Assessment and learning**. London: Sage, 2006a. p. 9-25.

BLOOM, B. S. **Taxonomy of educational objectives**: The classification of educational goals. New York: Longmans, 1969.

BLOOM, B. **Taxonomy of educational objectives**. New York: David McKay, 1956.

BLOOM, B.; HASTINGS, J.; MADAUS, G. **Handbook of formative and summative evaluation of student learning**. New York: MacGraw Hill, 1971.

BOUD, D.; FALCHIKOV, N. Aligning assessment with long-term learning. **Assessment & Evaluation in Higher Education**, v. 31, n. 4, p. 399-413, 2006.

BOUD, D.; FALCHIKOV, N. Rethinking assessment in higher education. In: BOUD, D.; FALCHIKOV, N. (Eds.). **Assessment and evaluation in higher education**. London: Routledge, 2007. p. 191-207.

BOWEN, R. S.; COOPER, M. M. Grading on a Curve as a Systemic Issue of Equity in Chemistry Education. **Journal of Chemical Education**, v. 99, n. 1, p. 185-194, 2021.

BRANDRIET, A. R. et al. Diagnosing changes in attitude in first-year college chemistry students with a shortened version of Bauer's semantic differential. **Chemistry Education Research and Practice**, v. 12, n. 2, p. 271-278, 2011.

BRANSFORD, J. D.; BROWN, A. L.; COCKING, R. R. **How people learn: brain, mind, experience, and school**. Washington, D.C.: National Academy Press, 2000.

BRASIL. **Base Nacional Comum Curricular**. Brasília: MEC, 2017.

BRASIL. **Guia de elaboração e revisão de itens**. INEP: Brasília, 2010. Disponível em: <<http://portal.inep.gov.br/guia-de-elaboracao-e-revisao-de-itens>> Acesso em 07 out. 2024.

BRASIL. **Lei de Diretrizes e Bases da Educação Nacional**. Brasília: MEC, 1996.

BROIETTI, F. C. D. **O ENEM, o vestibular e o ensino de Química: o caso da Universidade Estadual de Londrina**. Centro de Ciências Exatas, 2021. (tese de doutorado). Universidade Estadual de Londrina.

BROIETTI, F. C. D.; NORA, P. S.; COSTA, S. L. R. Dimensions of science learning: a study on PISA test questions involving chemistry content. **Revista Acta Scientiae**, v. 21, p. 95, 2019.

BROIETTI, F. C. D.; PASSOS, M. M.; SANTIN FILHO, O.; SOUZA, J. N. Alguns significados da expressão “deslocar o equilíbrio” em formandos do curso de Licenciatura em Química. **Ensaio: Pesquisa em Educação em Ciências** (Impresso), v. 15, p. 217-233, 2013.

BROIETTI, F. C. D.; SANTIN FILHO, O.; PASSOS, M. M. Uma análise da temática: avaliação em química em artigos de revistas da área de ensino no Brasil. **Revista Brasileira de Ensino de Química**, v. 10, n. 2, p. 42-54, 2015.

BROMAN, K.; PARCHMANN, I. Students' application of chemical concepts when solving chemistry problems in different contexts. **Chemistry Education Research and Practice**, v. 15, n. 4, p. 516-529, 2014.

BURROWS, N. L.; MOORING, S. R. Using concept mapping to uncover students' knowledge structures of chemical bonding concepts. **Chemistry Education Research and Practice**, v. 16, n. 1, p. 53-66, 2015.

BUSTOS-WORKS, C. *et al.* Moving toward inclusivity in chemistry by developing data-based instructional tasks aimed at increasing students' self-perception as capable learners who belong in STEM. **Journal of Chemical Education**, v. 99, n. 1, p. 177-184, 2021.

CABRAL, T. L. G.; ASSAI, N. D. S.; COSTA, S. L. R. Faraday's laws and the flipped classroom: a proposal for distance learning. **Research, Society and Development**, v. 10, 2021.

CANELAS, D. A.; HILL, J. L.; NOVICKI, A. Cooperative learning in organic chemistry increases student assessment of learning gains in key transferable skills. **Chemistry Education Research and Practice**, v. 18, n. 3, p. 441-456, 2017.

CARLE, M. S.; FLYNN, A. B. Essential learning outcomes for delocalization (resonance) concepts: how are they taught, practiced, and assessed in organic chemistry?. **Chemistry Education Research and Practice**, v. 21, n. 2, p. 622-637, 2020.

CARROLL, M. K. Moving from Recommendations to Innovations: Increasing the Relevancy and Effectiveness of Chemistry Education. **Journal of Chemical Education**, v. 90, n. 7, p. 816-819, 2013.

CHEE, Y. S.; TAN, K. C. D. Becoming chemists through game-based inquiry learning: The case of legends of Alkhimia. **Chemistry Education Research and Practice**, v. 13, n. 1, p. 10-22, 2012.

CHEN, B.; WEI, B. Examining chemistry teachers' use of curriculum materials: in view of teachers' pedagogical content knowledge. **Chemistry Education Research and Practice**, v. 16, n. 2, p. 260-272, 2015.

CHEN, K. *et al.* The individual experience of online chemistry teacher education in China: Coping with COVID-19 pandemic. **Journal of Chemical Education**, v. 97, n. 9, p. 3265-3270, 2020.

CHEN, Y. C.; WILSON, K.; LIN, H. S. Identifying the challenging characteristics of systems thinking encountered by undergraduate students in chemistry problem-solving of gas laws. **Chemistry Education Research and Practice**, v. 20, n. 3, p. 594-605, 2019.

CHEUNG, D. Using diagnostic assessment to help teachers understand the chemistry of the lead-acid battery. **Chemistry Education Research and Practice**, v. 12, n. 2, p. 228-237, 2011.

CHIANG, K. H. Relationship between research and teaching in doctoral education in UK universities. **Higher Education Policy**, v. 17, n. 1, p. 71-88, 2004.

CHILDS, P. E. Improving chemical education: turning research into effective practice. **Chemistry Education Research and Practice**, v. 10, n. 3, p. 189-203, 2009.

CHRISTIAN, B. N.; YEZIERSKI, E. J. Development and validation of an instrument to measure student knowledge gains for chemical and physical change for grades 6–8. **Chemistry Education Research and Practice**, v. 13, n. 3, p. 384-393, 2012.

ÇİÇEK, Ö.; ILHAN, N. A. İ. L. Evaluating interest in acids–bases: development of an acid–base interest scale (ABIS) and assessment of pre-service science teachers' interest. **Chemistry Education Research and Practice**, v. 18, n. 4, p. 630-640, 2017.

CLARIVATE ANALYTICS. 2023 **Journal Citation Reports**. 2024.

COLLINS III, John W.; O'BRIEN, Nancy Patricia. **The Greenwood Dictionary of Education**. Westport: Greenwood Press, 2003.

COLLINS, J. S.; OLESIK, S. V. The important role of chemistry department chairs and recommendations for actions they can enact to advance black student success. **Journal of Chemical Education**, v. 98, n. 7, p. 2209-2220, 2021.

COOPER, A. K.; OLIVER-HOYO, M. T. Argument construction in understanding noncovalent interactions: a comparison of two argumentation frameworks. **Chemistry Education Research and Practice**, v. 17, n. 4, p. 1006-1018, 2016.

COOPER, M. M. *et al.* Reliable multi method assessment of metacognition use in chemistry problem solving. **Chemistry Education Research and Practice**, v. 9, n. 1, p. 18-24, 2008.

COSTA, S. L. R. **Práticas científicas no ensino de ciências: características, compreensões e contextos das publicações**. 2021. 107 f. Dissertação (Mestrado em Ensino de Ciências e Educação Matemática) – Universidade Estadual de Londrina, Londrina, 2021.

COSTA, S. L. R.; BORTOLOCI, N. B.; BROIETTI, F. C. D.; VIEIRA, R. M.; TENREIRO-VIEIRA, C. Pensamento crítico no ensino de ciências e educação matemática: uma revisão bibliográfica sistemática. **Investigações em Ensino de Ciências** (Online), v. 26, p. 145-168, 2021.

COSTA, S. L. R.; BROIETTI, F. C. D. Potencialidades do microensino reflexivo na formação inicial de professores de Química. **Revista Docência do Ensino Superior**, v. 11, p. 1-19, 2021.

COSTA, S. L. R.; BROIETTI, F. C. D. Scientific practices in science education: a systematic review of research characteristics and trends. **Ensino & Pesquisa**, v. 19, p. 203-219, 2021.

COSTA, S. L. R.; BROIETTI, F. C. D. Scientific practices in science education publications: an analysis of research contexts. **Science Education International**, v. 32, p. 282-291, 2021.

COSTA, S. L. R.; BROIETTI, F. C. D. Towards greater understandings of scientific practices in science education: an analysis of the publications. **Revista Acta Scientiae**, v. 24, p. 151-182, 2022.

COSTA, S. L. R.; BROIETTI, F. C. D. What are scientific practices and why are they relevant in science education? **Revista Brasileira de Ensino de Ciência e Tecnologia**, v. 16, p. 1-22, 2023.

COSTA, S. L. R.; BROIETTI, F. C. D.; OBARA, Cássia Emi. Identifying scientific practices in a Science, Technology and Society themed workshop. **Acta Didactica Napocensia**, v. 14, p. 181-193, 2021.

COSTA, S. L. R.; BROIETTI, F. C. D.; PASSOS, M. M. The levels and nature of pre-service chemistry teachers' reflections in a public university in southern Brazil. **Problems of Education in the Twenty First Century**, v. 78, p. 147-166, 2020.

COSTA, S. L. R.; OBARA, C. E.; BROIETTI, F. C. D. Critical thinking in science education and mathematics education: research trends of 2010-2019. **Research, Society and Development**, v. 9, 2020.

COSTA, S. L. R.; OBARA, Cássia Emi; BROIETTI, F. C. D. Critical thinking in science education publications: the research contexts. **International Journal of Development Research**, v. 10, p. 39438-39448, 2020.

COX Jr, C. T. *et al.* Advancements in curriculum and assessment by the use of IMMEX technology in the organic laboratory. **Chemistry Education Research and Practice**, v. 9, n. 2, p. 163-168, 2008.

CRUJEIRAS-PÉREZ, B.; BROCOS, P. Pre-service teachers' use of epistemic criteria in the assessment of scientific procedures for identifying microplastics in beach sand. **Chemistry Education Research and Practice**, v. 22, n. 2, p. 237-246, 2021.

DALGARNO, B. *et al.* Effectiveness of a virtual laboratory as a preparatory resource for distance education chemistry students. **Computers & Education**, v. 53, n. 3, p. 853-865, 2009.

DANCZAK, S. M.; THOMPSON, C. D.; OVERTON, T. L. Development and validation of an instrument to measure undergraduate chemistry students' critical thinking skills. **Chemistry Education Research and Practice**, v. 21, n. 1, p. 62-78, 2020.

DANGUR, V. *et al.* Learning quantum chemistry via a visual-conceptual approach: students' bidirectional textual and visual understanding. **Chemistry Education Research and Practice**, v. 15, n. 3, p. 297-310, 2014.

DAVENPORT, J. L.; RAFFERTY, A. N.; YARON, D. J. Whether and how authentic contexts using a virtual chemistry lab support learning. **Journal of Chemical Education**, v. 95, n. 8, p. 1250-1259, 2018.

DEMIRDÖĞEN, B.; UZUNTIRYAKI-KONDAKÇI, E. Closing the gap between beliefs and practice: Change of pre-service chemistry teachers' orientations during a PCK-based NOS course. **Chemistry Education Research and Practice**, v. 17, n. 4, p. 818-841, 2016.

DENG, J. M.; FLYNN, A. B. Reasoning, granularity, and comparisons in students' arguments on two organic chemistry items. **Chemistry Education Research and Practice**, v. 22, n. 3, p. 749-771, 2021.

DENG, Y.; WANG, H. Research on evaluation of Chinese students' competence in written scientific argumentation in the context of chemistry. **Chemistry Education Research and Practice**, v. 18, n. 1, p. 127-150, 2017.

DERMAN, A.; EILKS, I. Using a word association test for the assessment of high school students' cognitive structures on dissolution. **Chemistry Education Research and Practice**, v. 17, n. 4, p. 902-913, 2016.

DING, N.; HARSKAMP, E. G. Collaboration and peer tutoring in chemistry laboratory education. **International Journal of Science Education**, v. 33, n. 6, p. 839-863, 2011.

DONNELLY, J.; HERNÁNDEZ, F. E. Fusing a reversed and informal learning scheme and space: student perceptions of active learning in physical chemistry. **Chemistry Education Research and Practice**, v. 19, n. 2, p. 520-532, 2018.

DOOD, A. J. *et al.* Analyzing explanations of substitution reactions using lexical analysis and logistic regression techniques. **Chemistry Education Research and Practice**, v. 21, n. 1, p. 267-286, 2020.

DORI, Y. J.; SASSON, I. Assessment in the context of learning: a multidimensional approach. In: **Assessment in Education: Principles, Policies and Practices**, v. 15, n. 2, p. 183-196, 2008.

DRESSEL, H. F.; MAYHEW, L. B. Toward a rational approach to educational assessment. **Educational Leadership**, v. 12, n. 3, p. 195-200, 1954.

DRESSEL, P. L.; MAYHEW, L. B. **General education**: explorations in evaluation. Washington, DC: American Council on Education, 1954.

EARL, L. M. **Assessment as learning: using classroom assessment to maximize student learning**. Corwin Press, 2012.

EATON, A. C. *et al.* Situating sustainable development within secondary chemistry education via systems thinking: a depth study approach. **Journal of Chemical Education**, v. 96, n. 12, p. 2968-2974, 2019.

EKIZ-KIRAN, B.; BOZ, Y.; OZTAY, E. S. Development of pre-service teachers' pedagogical content knowledge through a PCK-based school experience course. **Chemistry Education Research and Practice**, v. 22, n. 2, p. 415-430, 2021.

EL-FARARGY, N. Evaluation of a chemistry curriculum intervention using the Perry model of intellectual development. **Chemistry Education Research and Practice**, v. 11, n. 2, p. 98-106, 2010.

EL-KHAWAS, E. Defining the role of academics in accountability. **Higher Education Management and Policy**, v. 21, n. 1, p. 1-12, 2009.

ELMAS, R. *et al.* The intellectual demands of the intended chemistry curriculum in Czechia, Finland, and Turkey: a comparative analysis based on the revised Bloom's taxonomy. **Chemistry Education Research and Practice**, v. 21, n. 3, p. 839-851, 2020.

ELMGREN, M. *et al.* Comparison and evaluation of learning outcomes from an international perspective: Development of a best-practice process. **Journal of Chemical Education**, v. 92, n. 3, p. 427-432, 2015.

EMENIKE, M. E. *et al.* Results from a national needs assessment survey: A view of assessment efforts within chemistry departments. **Journal of Chemical Education**, v. 90, n. 5, p. 561-567, 2013.

ERDMANN, M. A.; MARCH, J. L. Video reports as a novel alternate assessment in the undergraduate chemistry laboratory. **Chemistry Education Research and Practice**, v. 15, n. 4, p. 650-657, 2014.

FAUTCH, J. M. The flipped classroom for teaching organic chemistry in small classes: is it effective? **Chemistry Education Research and Practice**, v. 16, n. 1, p. 179-186, 2015.

FERNANDES, D. Avaliar para melhorar as aprendizagens: análise e discussão de algumas questões essenciais. In: **Turma Mais e sucesso escolar**: contributos teóricos e práticos. 2011. p. 81-107.

FERNANDES, D. **Para uma fundamentação e melhoria das práticas de avaliação pedagógica**. Projeto de Monitorização e Investigação em Avaliação Pedagógica. Lisboa: Universidade de Lisboa; Instituto de Educação; Direção-Geral da Educação, 2020.

FERREIRA, A. C; BROIETTI, F. C. D. Avaliação no ensino de química: uma revisão em anais da RASBQ. In: **Anais do IV Congresso Paranaense de Educação Química**. ISSN – 2238-295X. Universidade Federal do Paraná – Curitiba/PR. p. 330 – 338. 26, 27 e 28 agosto 2015.

FIGUEIREDO, A. **A avaliação na educação**: práticas e desafios. Porto: Porto Editora, 2014.

FINK, A. **Conducting research literature reviews**: From the Internet to paper. 2 ed. Thousand Oaks, Estados Unidos: Sage, 2005.

FLYNN, A. B. Structure and evaluation of flipped chemistry courses: organic & spectroscopy, large and small, first to third year, English and French. **Chemistry Education Research and Practice**, v. 16, n. 2, p. 198-211, 2015.

FLYNN, A. B.; FEATHERSTONE, R. B. Language of mechanisms: exam analysis reveals students' strengths, strategies, and errors when using the electron-pushing formalism (curved arrows) in new reactions. **Chemistry Education Research and Practice**, v. 18, n. 1, p. 64-77, 2017.

FOX, L. J.; ROEHRIG, G. H. Nationwide survey of the undergraduate physical chemistry course. **Journal of Chemical Education**, v. 92, n. 9, p. 1456-1465, 2015.

FREITAS, M. Avaliação e formação de professores: desafios e possibilidades. *Revista Brasileira de Educação*, v. 24, n. 75, p. 1-20, 2019.

FULMER, K.; SULLIVAN, D. M.; HAYES, J. R. Promoting assessment literacy among educators. **Journal of Educational Assessment**, v. 25, n. 3, p. 221-239, 2018.

GALLOWAY, K. R. et al. Patterns of reactions: a card sort task to investigate students' organization of organic chemistry reactions. **Chemistry Education Research and Practice**, v. 20, n. 1, p. 30-52, 2019.

GENCER, S.; AKKUS, H. The topic-specific nature of experienced chemistry teachers' pedagogical content knowledge in the topics of interactions between chemical species and states of matter. **Chemistry Education Research and Practice**, v. 22, n. 2, p. 498-512, 2021.

GHANI, I. A. et al. Enhancing students' HOTS in laboratory educational activity by using concept map as an alternative assessment tool. **Chemistry Education Research and Practice**, v. 18, n. 4, p. 849-874, 2017.

GIBBONS, R. E. et al. Assessment in postsecondary chemistry education: a comparison of course types. **Assessment Update**, v. 30, n. 3, p. 1-16, 2018.

GIBBS, G.; SIMPSON, C. Conditions under which assessment supports students' learning. **Journal of Learning Sciences**, v. 13, n. 3, p. 313-331, 2004.

GILEWSKI, A. et al. Does linking help? Effects and student perceptions of a learner-centered assessment implemented in introductory chemistry. **Chemistry Education Research and Practice**, v. 20, n. 2, p. 399-411, 2019.

GLASER, R. Education and thinking: the role of knowledge. **American Psychologist**, v. 39, n. 2, p. 93-104, 1984.

GOK, T.; GOK, O. Peer instruction in chemistry education: Assessment of students' learning strategies, conceptual learning and problem solving. In **Asia-Pacific Forum on Science Learning & Teaching**, v. 17, n. 1, 2016.

GRAULICH, N. Intuitive judgments govern students' answering patterns in multiple-choice exercises in organic chemistry. **Journal of Chemical Education**, v. 92, n. 2, p. 205-211, 2015.

GÜLTEPE, N. Pre-service Chemistry Teachers' Understanding About Equilibria in Acid-Base Solutions. **Pedagogical Research**, v. 6, n. 4, 2021.

HADJI, C. **Avaliação desmistificada**. Porto Alegre, Artmed Editora. 2001.

HAHN, U.; CHATER, N. Representation and mapping in probabilistic reasoning. **Cognitive Science**, v. 22, n. 4, p. 651-687, 1998.

HANSON, R. Designing e-content for teaching basic analytical chemistry in higher education: a baseline study. **Science Education International**, v. 31, n. 1, p. 22-28, 2020.

HANSON, R. Employing Microscience Equipment to Promote Chemistry Education Through Constructivist Hands-and Minds-on Activities. **Interchange**, v. 52, n. 4, p. 521-544, 2021.

HARLEN, W. On the relationship between assessment for formative and summative purposes. In: GARDNER, J. (Ed.). **Assessment and learning**. London: Sage, 2006. p. 103-118.

HARLEN, W. Teachers' summative practices and assessment for learning: tensions and synergies. **Curriculum Journal**, v. 16, n. 2, p. 207-223, 2005.

HARSH, J. A. Designing performance-based measures to assess the scientific thinking skills of chemistry undergraduate researchers. **Chemistry Education Research and Practice**, v. 17, n. 4, p. 808-817, 2016.

HARSH, J.; ESTEB, J. J.; MALTese, A. V. Evaluating the development of chemistry undergraduate researchers' scientific thinking skills using performance-data: first findings from the performance assessment of undergraduate research (PURE) instrument. **Chemistry Education Research and Practice**, v. 18, n. 3, p. 472-485, 2017.

HARSHMAN, J.; YEZIERSKI, E. Characterizing high school chemistry teachers' use of assessment data via latent class analysis. **Chemistry Education Research and Practice**, v. 17, n. 2, p. 296-308, 2016.

HARSHMAN, J.; YEZIERSKI, E. Guiding teaching with assessments: high school chemistry teachers' use of data-driven inquiry. **Chemistry Education Research and Practice**, v. 16, n. 1, p. 93-103, 2015.

HARSHMAN, J.; YEZIERSKI, E. Test-retest reliability of the adaptive chemistry assessment survey for teachers: Measurement error and alternatives to correlation. **Journal of Chemical Education**, v. 93, n. 2, p. 239-247, 2016.

HART, C. **Doing a literature review**: releasing the social science research imagination. London: SAGE Publications, 1998.

HATTIE, J. **Visible Learning**: A Synthesis of Over 800 Meta-Analyses Relating to Achievement. Routledge, 2009.

HATTIE, J.; TIMPERLEY, H. The power of feedback. **Review of Educational Research**, v. 77, n. 1, p. 81-112, 2007.

HENNAH, N. A novel practical pedagogy for terminal assessment. **Chemistry Education Research and Practice**, v. 20, n. 1, p. 95-106, 2019.

HERRIDGE, M. et al. Variation in chemistry instructors' evaluations of student written responses and its impact on grading. **Chemistry Education Research and Practice**, v. 22, n. 4, p. 948-972, 2021.

HERRMANN-ABELL, C. F.; DEBOER, G. E. Using distractor-driven standards-based multiple-choice assessments and Rasch modeling to investigate hierarchies of chemistry misconceptions and detect structural problems with individual items. **Chemistry Education Research and Practice**, v. 12, n. 2, p. 184-192, 2011.

HIDI, S.; RENNINGER, K. A. The four phase model of interest development. **Educational Psychologist**, v. 41, n. 2, p. 111-127, 2006.

HOLME, T. *et al.* Enhancing the role of assessment in curriculum reform in chemistry. **Chemistry Education Research and Practice**, v. 11, n. 2, p. 92-97, 2010.

HOLME, T.; MURPHY, K. Assessing conceptual and algorithmic knowledge in general chemistry with ACS exams. **Journal of Chemical Education**, v. 88, n. 9, p. 1217-1222, 2011.

HORN, S.; HERNICK, M. Improving student understanding of lipids concepts in a biochemistry course using test-enhanced learning. **Chemistry Education Research and Practice**, v. 16, n. 4, p. 918-928, 2015.

HOUSEKNECHT, J. B. et al. Effectiveness of the active learning in organic chemistry faculty development workshops. **Chemistry Education Research and Practice**, v. 21, n. 1, p. 387-398, 2020.

HRIN, T.; MILENKOVIĆ, D.; SEGEDINAC, M. Diagnosing the quality of high school students' and pre-service chemistry teachers' cognitive structures in organic chemistry by using students' generated systemic synthesis questions. **Chemistry Education Research and Practice**, v. 19, n. 1, p. 305-318, 2018.

HUANG, J. Successes and challenges: online teaching and learning of chemistry in higher education in China in the time of COVID-19. **Journal of Chemical Education**, v. 97, n. 9, p. 2810-2814, 2020.

INEP. **Sistema de Avaliação da Educação Básica**. Brasília: INEP, 2020.

IRBY, S. M. et al. Use of a card sort task to assess students' ability to coordinate three levels of representation in chemistry. **Chemistry Education Research and Practice**, v. 17, n. 2, p. 337-352, 2016.

JAGODZIŃSKI, P.; WOLSKI, R. Assessment of application technology of natural user interfaces in the creation of a virtual chemical laboratory. **Journal of Science Education and Technology**, v. 24, n. 1, p. 16-28, 2015.

JOHNSTON, P. R.; WATTERS, D. J.; BROWN, C. L.; LOUGHLIN, W. A. An investigation into student perceptions towards mathematics and their performance in first year chemistry: introduction of online maths skills support. **Chemistry Education Research and Practice**, v. 17, n. 4, p. 1203-1214, 2016.

KARARO, A. T. et al. Predictions and constructing explanations: an investigation into introductory chemistry students' understanding of structure–property relationships. **Chemistry Education Research and Practice**, v. 20, n. 1, p. 316-328, 2019.

KARTOWAGIRAN, B.; ROHAETI, E. A critical thinking assessment model integrated with science process skills on chemistry for senior high school. **European Journal of Educational Research**, v. 10, n. 1, p. 285-298, 2021.

KIM, T.; WRIGHT, L. K.; MILLER, K. An examination of students' perceptions of the Kekulé resonance representation using a perceptual learning theory lens. **Chemistry Education Research and Practice**, v. 20, n. 4, p. 659-666, 2019.

KINGIR, S.; GEBAN, O.; GUNEL, M. How does the science writing heuristic approach affect students' performances of different academic achievement levels? A case for high school chemistry. **Chemistry Education Research and Practice**, v. 13, n. 4, p. 428-436, 2012.

KISHBAUGH, T. L. et al. Measuring beyond content: a rubric bank for assessing skills in authentic research assignments in the sciences. **Chemistry Education Research and Practice**, v. 13, n. 3, p. 268-276, 2012.

KOEPER, I. et al. Turning chemistry education on its head: Design, experience and evaluation of a learning-centred 'Modern Chemistry' subject. **Journal of University Teaching & Learning Practice**, v. 17, n. 3, p. 13, 2020.

KOUSA, P. et al. Low-achieving students' attitudes towards learning chemistry and chemistry teaching methods. **Chemistry Education Research and Practice**, v. 19, n. 2, p. 431-441, 2018.

KOUTALIDI, S.; SCOULLOS, M. Biogeochemical cycles for combining chemical knowledge and ESD issues in Greek secondary schools part I: designing the didactic materials. **Chemistry Education Research and Practice**, v. 17, n. 1, p. 10-23, 2016.

KRAPP, A. Basic needs and the development of interest and intrinsic motivational orientations. **Learning and Instruction**, v. 15, n. 5, p. 381-395, 2005.

KRAPP, A. Structural and dynamic aspects of interest development: theoretical considerations from an ontogenetic perspective. **Learning and Instruction**, v. 12, n. 4, p. 383-409, 2002.

KRAUSE, M. et al. On the development and assessment of a computer-based learning and assessment environment for the transition from lower to upper secondary chemistry education. **Chemistry Education Research and Practice**, v. 14, n. 3, p. 345-353, 2013.

KUH, G. D.; KINZIE, J.; BUCKLEY, J. A.; BRIDGES, B.; HAYEK, J. C. **What matters to student success**: A review of the literature. National Postsecondary Education Cooperative, 2005.

KUHN, D. A developmental model of critical thinking. **Educational Research**, v. 28, n. 2, p. 16-26, 1999.

LAMICHHANE, R.; RECK, C.; MALTESE, A. V. Undergraduate chemistry students' misconceptions about reaction coordinate diagrams. **Chemistry Education Research and Practice**, v. 19, n. 3, p. 834-845, 2018.

LAU, P. N. Enhancing formative and self-assessment with video playback to improve critique skills in a titration laboratory. **Chemistry Education Research and Practice**, v. 21, n. 1, p. 178-188, 2020.

LAWRIE, G. A. et al. Personal journeys of teachers: an investigation of the development of teacher professional knowledge and skill by expert tertiary chemistry teachers. **Chemistry Education Research and Practice**, v. 20, n. 1, p. 132-145, 2019.

LAZENBY, K. et al. From ideas to items: a primer on the development of ordered multiple-choice items for investigating the progression of learning in higher education STEM. **Journal of Chemical Education**, v. 98, n. 3, p. 714-729, 2021.

LEVY, S. T.; WILENSKY, U. Students' learning with the connected chemistry (CC1) curriculum: Navigating the complexities of the particulate world. **Journal of Science Education and Technology**, v. 18, n. 3, p. 243-254, 2009.

LEWIS, S. E. Examining evidence for external and consequential validity of the first term general chemistry exam from the ACS Examinations Institute. **Journal of Chemical Education**, v. 91, n. 6, p. 793-799, 2014.

LEWIS, S. E.; SHAW, J. L.; FREEMAN, K. A. Establishing open-ended assessments: investigating the validity of creative exercises. **Chemistry Education Research and Practice**, v. 12, n. 2, p. 158-166, 2011.

LOERTSCHER, J. et al. Development and use of a construct map framework to support teaching and assessment of noncovalent interactions in a biochemical context. **Chemistry Education Research and Practice**, v. 19, n. 4, p. 1151-1165, 2018.

LOPES, A. C.; MACEDO, E. **Teorias de currículo**. Cortez Editora, 2014.

LOPEZ, E. et al. Validating the use of concept-mapping as a diagnostic assessment tool in organic chemistry: implications for teaching. **Chemistry Education Research and Practice**, v. 12, n. 2, p. 133-141, 2011.

LOZANO, R.; WATSON, M. K. Chemistry education for sustainability: Assessing the chemistry curricula at Cardiff University. **Educación Química**, v. 24, n. 2, p. 184-192, 2013.

LU, H.; JIANG, Y.; BI, H. Development of a measurement instrument to assess students' proficiency levels regarding galvanic cells. **Chemistry Education Research and Practice**, v. 21, n. 2, p. 655-667, 2020.

MA, J. Getting a bigger picture in less time: viewing curriculum reform in a Chinese graduate chemistry program through the lens of an organic structure analysis course. **Journal of Chemical Education**, v. 88, n. 12, p. 1639-1643, 2011.

MACK, M. R. et al. Metrics and methods used to compare student performance data in chemistry education research articles. **Journal of Chemical Education**, v. 96, n. 3, p. 401-413, 2019.

MAGNUSSON, S.; KRAKOVSKY, M.; PALMER, D. A. Conceptualizing pedagogical content knowledge in science education. In: **Teaching Science: A Handbook for Teachers**. New York: Routledge, 1999. p. 17-35.

MARKHAM, K.; BROWN, S.; HATFIELD, D. **Learning through inquiry: a guide for science teachers**. Washington, D.C.: National Science Teachers Association, 1994.

MARTIN, J. R.; GARDNER, J.; HOOD, P. **Teaching and learning science through inquiry**. 2. ed. London: Routledge, 2000.

MARTIN, S. N.; SCANTLEBURY, K. More than a conversation: Using cogenerative dialogues in the professional development of high school chemistry teachers. **Educational Assessment, Evaluation and Accountability**, v. 21, n. 2, p. 119-136, 2009.

MATHABATHE, K. C.; POTGIETER, M. Metacognitive monitoring and learning gain in foundation chemistry. **Chemistry Education Research and Practice**, v. 15, n. 1, p. 94-104, 2014.

MCMANUS, S. **Formative assessment for teachers and students**. Council of Chief State School Officers, 2008.

MEYER, D. K.; TURNER, J. C. Discovering emotion in classroom management: a qualitative analysis of middle school teachers' beliefs. **Educational Administration Quarterly**, v. 42, n. 1, p. 9-20, 2006.

MOON, A. et al. Application and testing of a framework for characterizing the quality of scientific reasoning in chemistry students' writing on ocean acidification. **Chemistry Education Research and Practice**, v. 20, n. 3, p. 484-494, 2019.

MORAIS, C.; ARAÚJO, J. L. Awakening to chemistry through storytelling and practical activities: middle school students interacting with pre-school children. **Chemistry Education Research and Practice**, v. 20, n. 1, p. 302-315, 2019.

MORRISON, D.; ROSS, S. M.; KEMP, J. E. **Educational Technology: A Systematic Approach to Design, Development, and Implementation**. Pearson, 2017.

MURCIANO-CALLES, J. Use of Kahoot for assessment in chemistry education: a comparative study. **Journal of Chemical Education**, v. 97, n. 11, p. 4209-4213, 2020.

NAGEL, M. L.; LINDSEY, B. A. Implementation of reasoning chain construction tasks to support student explanations in general chemistry. **Journal of Chemical Education**, v. 99, n. 2, p. 839-850, 2021.

NAKIBOGLU, C. Using word associations for assessing non-major science students' knowledge structure before and after general chemistry instruction: the case of atomic structure. **Chemistry Education Research and Practice**, v. 9, n. 4, p. 309-322, 2008.

NATIONAL RESEARCH COUNCIL. **Knowing what students know: the science and design of educational assessment**. Washington, D.C.: National Academies Press, 2001.

NEILES, K. Y.; TODD, I.; BUNCE, D. M. Establishing the validity of using network analysis software for measuring students' mental storage of chemistry concepts. **Journal of Chemical Education**, v. 93, n. 5, p. 821-831, 2016.

NICOL, D. J.; MACFARLANE-DICK, D. Formative assessment and self-regulated learning: a model and seven principles of good feedback practice. **Studies in Higher Education**, v. 31, n. 2, p. 199-218, 2006.

NIELSEN, W. et al. Pharmacology students' perceptions of creating multimodal digital explanations. **Chemistry Education Research and Practice**, v. 18, n. 2, p. 329-339, 2017.

NOTTIS, K. E.; KASTNER, M. E. The effect of instructional modality and prior knowledge on learning point group symmetry. **Journal of Science Education and Technology**, v. 14, n. 1, p. 51-58, 2005.

NOVAK, J. D.; GOWIN, D. B. **Learning how to learn**. New York: Cambridge University Press, 1984.

NYACHWAYA, J. M. et al. College chemistry students' use of memorized algorithms in chemical reactions. **Chemistry Education Research and Practice**, v. 15, n. 1, p. 81-93, 2014.

NYACHWAYA, J. M. et al. The development of an open-ended drawing tool: an alternative diagnostic tool for assessing students' understanding of the particulate nature of matter. **Chemistry Education Research and Practice**, v. 12, n. 2, p. 121-132, 2011.

O'DWYER, A.; CHILDS, P. Organic chemistry in action! Developing an intervention program for introductory organic chemistry to improve learners' understanding, interest, and attitudes. **Journal of Chemical Education**, v. 91, n. 7, p. 987-993, 2014.

OKOLI, C.; SCHABRAM, K. A guide to conducting a systematic literature review of information systems research. **Sprouts: Working Papers on Information Systems**, v. 9, n. 37, 2015. Disponível em: <http://sprouts.aisnet.org/9-37>. Acesso em: 4 nov. 2024.

OLIVER-HOYO, M.; BABILONIA-ROSA, M. A. Promotion of spatial skills in chemistry and biochemistry education at the college level. **Journal of Chemical Education**, v. 94, n. 8, p. 996-1006, 2017.

ÖZALP, D.; KAHVECI, A. Diagnostic assessment of student misconceptions about the particulate nature of matter from ontological perspective. **Chemistry Education Research and Practice**, v. 16, n. 3, p. 619-639, 2015.

PAGE, M. F. et al. The effect of teaching the entire academic year of high school chemistry utilizing abstract reasoning. **Chemistry Education Research and Practice**, v. 19, n. 2, p. 500-507, 2018.

PANTIWAITI, Y. et al. Instructional model of natural science in junior high schools, Batu-Malang. **Journal of Education and Practice**, v. 8, n. 8, 2017.

PARK, M. et al. The effect of computer models as formative assessment on student understanding of the nature of models. **Chemistry Education Research and Practice**, v. 18, n. 4, p. 572-581, 2017.

PEREIRA, A. Novas perspectivas para a avaliação na educação. **Revista Brasileira de Estudos da Educação**, v. 23, n. 3, p. 401-416, 2018.

PEREIRA, C. **Avaliação e educação: um caminho de reflexão**. Porto Alegre: Editora Texto, 2009.

PETTERSON, M. N. et al. Eliciting student thinking about acid–base reactions via app and paper–pencil based problem solving. **Chemistry Education Research and Practice**, v. 21, n. 3, p. 878-892, 2020.

PETTICREW, M.; ROBERTS, H. **Systematic reviews in the social sciences: a practical guide**. Malden: Blackwell Publishing, 2006.

PINTO, A. M. O Processo de Bologna: uma análise crítica. **Revista de Educação e Pesquisa em Contabilidade**, v. 4, n. 1, p. 25-38, 2010.

POPHAM, W. J. **Transformative assessment**. Alexandria: ASCD, 2008.

POPOVA, M.; JONES, T. Chemistry instructors' intentions toward developing, teaching, and assessing student representational competence skills. **Chemistry Education Research and Practice**, v. 22, n. 3, p. 733-748, 2021.

PRINS, G. T. et al. Selection of authentic modelling practices as contexts for chemistry education. **International Journal of Science Education**, v. 30, n. 14, p. 1867-1890, 2008.

PROPHET, R. B.; VLAARDINGERBROEK, B. The relevance of secondary school chemistry education in Botswana: a cognitive development status perspective. **International Journal of Educational Development**, v. 23, n. 3, p. 275-289, 2003.

PUGLIESE, G. STEM education: um panorama e sua relação com a educação brasileira. **Currículo sem Fronteiras**, v. 20, n. 1, p. 209-232, 2020.

PULLEN, R. et al. Investigating the viability of a competency-based, qualitative laboratory assessment model in first-year undergraduate chemistry. **Chemistry Education Research and Practice**, v. 19, n. 2, p. 629-637, 2018.

RAJE, S.; STITZEL, S. Strategies for effective assessments while ensuring academic integrity in general chemistry courses during COVID-19. **Journal of Chemical Education**, v. 97, n. 9, p. 3436-3440, 2020.

RALPH, V. R.; LEWIS, S. E. An explanative basis for the differential performance of students with low math aptitude in general chemistry. **Chemistry Education Research and Practice**, v. 20, n. 3, p. 570-593, 2019.

RAVI, M.; PUENTE-URBINA, A.; VAN BOKHOVEN, J. A. Identifying opportunities to promote systems thinking in catalysis education. **Journal of Chemical Education**, v. 98, n. 5, p. 1583-1593, 2021.

REES, S. W.; KIND, V.; NEWTON, D. Can language focused activities improve understanding of chemical language in non-traditional students? **Chemistry Education Research and Practice**, v. 19, n. 3, p. 755-766, 2018.

ROCABADO, G. A.; KOMPERDA, R.; LEWIS, J. E.; BARBERA, J. Addressing diversity and inclusion through group comparisons: a primer on measurement invariance testing. **Chemistry Education Research and Practice**, v. 21, n. 3, p. 969-988, 2020.

RODRIGUEZ, J. M. G. et al. Covariational reasoning and mathematical narratives: Investigating students' understanding of graphs in chemical kinetics. **Chemistry Education Research and Practice**, v. 20, n. 1, p. 107-119, 2019.

ROOTMAN-LE GRANGE, I.; BLACKIE, M. A. Assessing assessment: in pursuit of meaningful learning. **Chemistry Education Research and Practice**, v. 19, n. 2, p. 484-490, 2018.

RUSMINI, R.; SUYONO, S.; AGUSTINI, R. Analysis of science process skills of chemical education students through self project based learning (SjBL) in the pandemic COVID 19 era. **JOTSE**, v. 11, n. 2, p. 371-387, 2021.

RYAN, B. J. Line up, line up: using technology to align and enhance peer learning and assessment in a student centred foundation organic chemistry module. **Chemistry Education Research and Practice**, v. 14, n. 3, p. 229-238, 2013.

RYOO, K.; TOUTKOUSHIAN, E.; BEDELL, K. Exploring different types of assessment items to measure linguistically diverse students' understanding of energy and matter in chemistry. **Chemistry Education Research and Practice**, v. 19, n. 1, p. 149-166, 2018.

SAITTA, E. K. H.; BOWDON, M. A.; GEIGER, C. L. Incorporating service-learning, technology, and research supportive teaching techniques into the university chemistry classroom. **Journal of Science Education and Technology**, v. 20, n. 6, p. 790-795, 2011.

SANABRIA-RÍOS, D.; BRETZ, S. L. Investigating the relationship between faculty cognitive expectations about learning chemistry and the construction of exam questions. **Chemistry Education Research and Practice**, v. 11, n. 3, p. 212-217, 2010.

SANTOS, M. **Avaliação e suas dimensões: uma abordagem crítica**. Lisboa: Edições 70, 2012.

SANTOS-DÍAZ, S.; TOWNS, M. H. An all-female graduate student organization participating in chemistry outreach: a case study characterizing leadership in the community of practice. **Chemistry Education Research and Practice**, v. 22, n. 2, p. 532-553, 2021.

SASSON, I.; DORI, Y. J. A three-attribute transfer skills framework—part II: applying and assessing the model in science education. **Chemistry Education Research and Practice**, v. 16, n. 1, p. 154-167, 2015.

SCHAFER, A. G. et al. Visualizing chemistry teachers' enacted assessment design practices to better understand barriers to "best practices". **Chemistry Education Research and Practice**, v. 22, n. 2, p. 457-475, 2021.

SCHAFER, A. G.; YEZIERSKI, E. J. Chemistry critical friendships: investigating chemistry-specific discourse within a domain-general discussion of best practices for inquiry assessments. **Chemistry Education Research and Practice**, v. 21, n. 1, p. 452-468, 2020.

SCHAFER, A. G.; YEZIERSKI, E. J. Investigating high school chemistry teachers' assessment item generation processes for a solubility lab. **Chemistry Education Research and Practice**, v. 22, n. 1, p. 214-225, 2021.

SCHERER, R. et al. Developing a computer-based assessment of complex problem solving in Chemistry. **International Journal of STEM Education**, v. 1, n. 1, p. 1-15, 2014.

SCHMIDT-MCCORMACK, J. A. et al. Analysis of the role of a writing-to-learn assignment in student understanding of organic acid–base concepts. **Chemistry Education Research and Practice**, v. 20, n. 2, p. 383-398, 2019.

SCHNEIDER, J. L.; RUDER, S. M.; BAUER, C. F. Student perceptions of immediate feedback testing in student centered chemistry classes. **Chemistry Education Research and Practice**, v. 19, n. 2, p. 442-451, 2018.

SCHUCHARD, J.; NELSON, D.; WILLIAMS, R. Evaluation of diagnostic tools that tertiary teachers can apply to profile their students' conceptions. **International Journal of Science Education**, v. 39, n. 5, p. 565-586, 2017.

SCOTT, D.; FIRTH, D. Using control charts early in the quantitative analysis laboratory curriculum. **Journal of Chemical Education**, v. 96, n. 5, p. 1037-1041, 2019.

SCOTT, F. J. A simulated peer-assessment approach to improving student performance in chemical calculations. **Chemistry Education Research and Practice**, v. 15, n. 4, p. 568-575, 2014.

SCRIVEN, M. The methodology of evaluation. In: TYLER, R. W. (Ed.). **Educational evaluation: New roles, new methods**. Berkeley: University of California, 1967. p. 9-30.

SEERY, M. K. et al. Developing laboratory skills by incorporating peer-review and digital badges. **Chemistry Education Research and Practice**, v. 18, n. 3, p. 403-419, 2017.

SEGEDINAC, M. D. et al. The development of research in the field of chemistry education at the University of Novi Sad since the breakup of the Socialist Federal Republic of Yugoslavia. **CEPS Journal**, v. 10, n. 1, p. 103-124, 2020.

SENDUR, G. An examination of pre-service chemistry teachers' meaningful understanding and learning difficulties about aromatic compounds using a systemic assessment questions diagram. **Chemistry Education Research and Practice**, v. 21, n. 1, p. 113-140, 2020.

SENOCAK, E.; BALOGLU, M. The adaptation and preliminary psychometric properties of the Derived Chemistry Anxiety Rating Scale. **Chemistry Education Research and Practice**, v. 15, n. 4, p. 800-806, 2014.

SEVIAN, H. et al. Use of representation mapping to capture abstraction in problem solving in different courses in chemistry. **Chemistry Education Research and Practice**, v. 16, n. 3, p. 429-446, 2015.

SEVIAN, H.; ROBINSON, W. E. Clickers promote learning in all kinds of classes: small and large, graduate and undergraduate, lecture and lab. **Journal of College Science Teaching**, v. 40, n. 3, 2011.

SHAH, L.; RODRIGUEZ, C. A.; BARTOLI, M.; RUSHTON, G. T. Analysing the impact of a discussion-oriented curriculum on first-year general chemistry students' conceptions of relative acidity. **Chemistry Education Research and Practice**, v. 19, n. 2, p. 543-557, 2018.

SHEPARD, L. A. Formative assessment: caveat emptor. **Educational Psychologist**, v. 43, n. 3, p. 197-204, 2008.

SILVA, J. L. P.; MORADILLO, E. F. D. Avaliação, ensino e aprendizagem de ciências. **Ensaio: Pesquisa em Educação em Ciências**, Belo Horizonte, v. 4, n. 1, p. 28-39, 2002.

SILVA, R. **Desafios da avaliação educacional no século XXI**. Coimbra: Editora Universitária, 2017.

SIWEYA, H. J.; LETSOALO, P. Formative assessment by first-year chemistry students as predictor of success in summative assessment at a South African university. **Chemistry Education Research and Practice**, v. 15, n. 4, p. 541-549, 2014.

SMITH, K. C.; VILLARREAL, S. Using animations in identifying general chemistry students' misconceptions and evaluating their knowledge transfer relating to particle position in physical changes. **Chemistry Education Research and Practice**, v. 16, n. 2, p. 273-282, 2015.

SOUTHAM, D. C. et al. The timing of an experiment in the laboratory program is crucial for the student laboratory experience: acylation of ferrocene as a case study. **Chemistry Education Research and Practice**, v. 14, n. 4, p. 476-484, 2013.

SOUZA, A. C.; BROIETTI, F. C. D. Análise em anais do ENPEC sobre a temática Avaliação em Química. **ACTIO: Docência em Ciências**, v. 2, p. 122-142, 2017.

SOUZA, A. C.; BROIETTI, F. C. D. Análise em anais do ENPEC sobre a temática avaliação em química. **ACTIO: Docência em Ciências**. Curitiba-PR, v. 2, nº 1, p. 122-142, jan./jul. 2017.

STIGGINS, R. From formative assessment to assessment for learning: a path to success in standards-based schools. **Phi Delta Kappan**, v. 87, n. 4, p. 324-328, 2005.

STOWE, D.; COOPER, P. **Models of learning and assessment in science education**. Cham: Springer, 2019.

STOWE, K.; COOPER, T. Assessing learning in higher education: a guide for educators. **Journal of University Teaching & Learning Practice**, v. 16, n. 5, 2019. Disponível em: <https://ro.uow.edu.au/jutlp/vol16/iss5/6>. Acesso em: 4 nov. 2024.

STOWE, R. L. et al. You are what you assess: The case for emphasizing chemistry on chemistry assessments. **Journal of Chemical Education**, v. 98, n. 8, p. 2490-2495, 2021.

STULL, A. T.; GAINER, M. J.; HEGARTY, M. Learning by enacting: the role of embodiment in chemistry education. **Learning and Instruction**, v. 55, p. 80-92, 2018.

SUPALO, C. A. *et al.* Examining the use of adaptive technologies to increase the hands-on participation of students with blindness or low vision in secondary-school chemistry and physics. **Chemistry Education Research and Practice**, v. 17, n. 4, p. 1174-1189, 2016.

TAAGEPERA, M. et al. Integrating symmetry in stereochemical analysis in introductory organic chemistry. **Chemistry Education Research and Practice**, v. 12, n. 3, p. 322-330, 2011.

TAL, M.; HERSCOVITZ, O.; DORI, Y. J. Assessing teachers' knowledge: incorporating context-based learning in chemistry. **Chemistry Education Research and Practice**, v. 22, n. 4, p. 1003-1019, 2021.

TALANQUER, V.; POLLARD, J. Let's teach how we think instead of what we know. **Chemistry Education Research and Practice**, v. 11, n. 2, p. 74–83, 2010.

TALBERT, L. E. *et al.* Revisiting the use of concept maps in a large enrollment general chemistry course: implementation and assessment. **Chemistry Education Research and Practice**, v. 21, n. 1, p. 37-50, 2020.

TAN, K. C. D.; TABER, K. S.; LIEW, Y. Q.; TEO, K. L. A. A web-based ionisation energy diagnostic instrument: exploiting the affordances of technology. **Chemistry Education Research and Practice**, v. 20, n. 2, p. 412-427, 2019.

TASHIRO, J. *et al.* Characterizing change in students' self-assessments of understanding when engaged in instructional activities. **Chemistry Education Research and Practice**, v. 22, n. 3, p. 662-682, 2021.

TENREIRO-VIEIRA, C.; VIEIRA, R. M. **Construindo práticas didático-pedagógicas promotoras da literacia científica e do pensamento crítico**. Madrid: OEI – Organização dos Estados Ibero-americanos, 2014. (Iberciencia, n. 2). Disponível em: <http://www.ibercienciaoei.org/doc2.pdf>. Acesso em: 4 nov. 2024.

TENREIRO-VIEIRA, C.; VIEIRA, R. M. Promover o pensamento crítico em ciências na escolaridade básica: propostas e desafios. **Revista Latinoamericana de Estudios Educativos**, v. 15, n. 15, p. 36-49, 2019.

THOMAS, G. A review of the literature on the role of assessment in the learning process. **Assessment in Education: Principles, Policies and Practices**, v. 7, n. 2, p. 143-158, 2000.

TOPPING, K. J. Peer assessment. **Theory into Practice**, v. 48, n. 1, p. 20-27, 2009.

TURNER, M. *et al.* Motivations for and barriers to the implementation of diagnostic assessment practices—a case study. **Chemistry Education Research and Practice**, v. 12, n. 2, p. 142-157, 2011.

UNDERSANDER, M. A. *et al.* Probing the question order effect while developing a chemistry concept inventory. **Chemistry Education Research and Practice**, v. 18, n. 1, p. 45-54, 2017.

VACHLIOTIS, T.; SALTA, K.; TZOUGRAKI, C. Developing Basic Systems Thinking Skills for Deeper Understanding of Chemistry Concepts in High School Students. **Thinking Skills and Creativity**, v. 41, p. 100881, 2021.

VENKATARAMAN, B. Visualization and interactivity in the teaching of chemistry to science and non-science students. **Chemistry Education Research and Practice**, v. 10, n. 1, p. 62-69, 2009.

VILLAFANE, S. M. *et al.* Uncovering students' incorrect ideas about foundational concepts for biochemistry. **Chemistry Education Research and Practice**, v. 12, n. 2, p. 210-218, 2011.

VILLAFANE, S. M.; LEWIS, J. E. Exploring a measure of science attitude for different groups of students enrolled in introductory college chemistry. **Chemistry Education Research and Practice**, v. 17, n. 4, p. 731-742, 2016.

VISHNUMOLAKALA, V. R. *et al.* Latent constructs of the students' assessment of their learning gains instrument following instruction in stereochemistry. **Chemistry Education Research and Practice**, v. 17, n. 2, p. 309-319, 2016.

VLADUŠIĆ, R.; BUCAT, R.; OŽIĆ, M. Understanding of words and symbols by chemistry university students in Croatia. **Chemistry Education Research and Practice**, v. 17, n. 3, p. 474-488, 2016.

VOGELZANG, J.; ADMIRAAL, W. F. Classroom action research on formative assessment in a context-based chemistry course. **Educational Action Research**, v. 25, n. 1, p. 155-166, 2017.

VOM BROCKE, J.; SIMON, J.; PISTOR, A.; REIMER, U.; LÖRCHER, I. A guideline for standardizing literature reviews. **Business & Information Systems Engineering**, v. 1, n. 1, p. 23-35, 2009.

WALKER, J. P. *et al.* Using the laboratory to engage all students in science practices. **Chemistry Education Research and Practice**, v. 17, n. 4, p. 1098-1113, 2016.

WALLY, M.; LEVINGER, E.; GRAINGER, D. Employing popular children's literature to teach elementary school chemistry. **Journal of Chemical Education**, v. 82, n. 10, p. 1489-1495, 2005.

WALPUSKI, M.; ROPUHL, M.; SUMFLETH, E. Students' knowledge about chemical reactions—development and analysis of standard-based test items. **Chemistry Education Research and Practice**, v. 12, n. 2, p. 174-183, 2011.

WANG, Y.; LEWIS, S. E. Analytical chemistry students' explanatory statements in the context of their corresponding lecture. **Chemistry Education Research and Practice**, v. 21, n. 4, p. 1183-1198, 2020.

WARFA, A. R. M.; ODOWA, N. Creative exercises (CEs) in the biochemistry domain: an analysis of students' linking of chemical and biochemical concepts. **Chemistry Education Research and Practice**, v. 16, n. 4, p. 747-757, 2015.

WATSON, S. W. *et al.* Increasing chemistry students' knowledge, confidence, and conceptual understanding of pH using a collaborative computer pH simulation. **Chemistry Education Research and Practice**, v. 21, n. 2, p. 528-535, 2020.

WEINRICH, M. L.; SEVIAN, H. Capturing students' abstraction while solving organic reaction mechanism problems across a semester. **Chemistry Education Research and Practice**, v. 18, n. 1, p. 169-190, 2017.

WHEELER, L. B.; MAENG, J. L.; WHITWORTH, B. A. Teaching assistants' perceptions of a training to support an inquiry-based general chemistry laboratory course. **Chemistry Education Research and Practice**, v. 16, n. 4, p. 824-842, 2015.

WILIAM, D. Assessment: learning communities. **The Curriculum Journal**, v. 22, n. 3, p. 335-348, 2011.

WILIAM, D.; THOMPSON, M. Integrating assessment with instruction: what will it take to make it work? **Educational Psychologist**, v. 43, n. 1, p. 1-11, 2008.

WILLIAMS, N. A. *et al.* Improving student achievement and satisfaction by adopting a blended learning approach to inorganic chemistry. **Chemistry Education Research and Practice**, v. 9, n. 1, p. 43-50, 2008.

WILLIAMSON, V. M. *et al.* Impact of basic arithmetic skills on success in first-semester general chemistry. **Chemistry Education Research and Practice**, v. 21, n. 1, p. 51-61, 2020.

WILSON, M.; SCALISE, K. Assessment to improve learning in higher education: The BEAR Assessment System. **Higher Education**, v. 52, n. 4, p. 635-663, 2006.

WREN, D.; BARBERA, J. Psychometric analysis of the thermochemistry concept inventory. **Chemistry Education Research and Practice**, v. 15, n. 3, p. 380-390, 2014.

XU, Y.; LEWIS, J. E. Development and validation of the Attitude toward the Subject of Chemistry Inventory (ASCIv2). **Journal of Chemical Education**, v. 88, n. 12, p. 1697-1702, 2011.

YAMAN, F.; AYAS, A. Assessing changes in high school students' conceptual understanding through concept maps before and after the computer-based predict–observe–explain (CB-POE) tasks on acid–base chemistry at the secondary level. **Chemistry Education Research and Practice**, v. 16, n. 4, p. 843-855, 2015.

YE, L. *et al.* The impact of coupling assessments on conceptual understanding and connection-making in chemical equilibrium and acid–base chemistry. **Chemistry Education Research and Practice**, v. 21, n. 3, p. 1000-1012, 2020.

YE, L.; LEWIS, S. E. Looking for links: examining student responses in creative exercises for evidence of linking chemistry concepts. **Chemistry Education Research and Practice**, v. 15, n. 4, p. 576-586, 2014.

YIK, B. J. *et al.* Development of a machine learning-based tool to evaluate correct Lewis acid–base model use in written responses to open-ended formative assessment items. **Chemistry Education Research and Practice**, v. 22, n. 4, p. 866-885, 2021.

YORK, S. *et al.* Applications of systems thinking in STEM education. **Journal of Chemical Education**, v. 96, n. 12, p. 2742-2751, 2019.

YOUNG, K. R. *et al.* Tired of failing students? Improving student learning using detailed and automated individualized feedback in a large introductory science course. **Innovative Higher Education**, v. 46, n. 2, p. 133-151, 2021.

YÜKSEL, M. Evaluating the effectiveness of the chemistry education by using the analytic hierarchy process. **International Education Studies**, v. 5, n. 5, p. 79-91, 2012.

YUSUF, M.; TAYLOR, P. C.; DAMANHURI, M. I. M. Designing critical pedagogy to counteract the hegemonic culture of the traditional chemistry classroom. **Issues in Educational Research**, v. 27, n. 1, p. 168-184, 2017.

ZANE, M.; TUCCI, V. K. Exploring the information literacy needs and values of high school chemistry teachers. **Journal of Chemical Education**, v. 93, n. 3, p. 406-412, 2016.

ZEMEL, Y. *et al.* Preservice teachers' enactment of formative assessment using rubrics in the inquiry-based chemistry laboratory. **Chemistry Education Research and Practice**, v. 22, n. 4, p. 1074-1092, 2021.

ZHANG, L.; WANG, L.; TREAGUST, D. F. Discipline-specific cognitive factors that influence grade 9 students' performance in chemistry. **Chemistry Education Research and Practice**, v. 22, n. 4, p. 813-841, 2021.

ZOLLER, U. Algorithmic, LOCS and HOCS (chemistry) exam questions: Performance and attitudes of college students. **International Journal of Science Education**, v. 24, n. 2, p. 185-203, 2002.

ZOLLER, U. Education in environmental chemistry: Setting the agenda and recommending action. A workshop report summary. **Journal of Chemical Education**, v. 82, n. 8, p. 1237, 2005.

APÊNDICES

APÊNDICE A

Codificação dos 205 artigos revistos nesse estudo

Código	Referência
A001	KARTOWAGIRAN, B.; ROHAETI, E. A Critical Thinking Assessment Model Integrated with Science Process Skills on Chemistry for Senior High School. European Journal of Educational Research , v. 10, n. 1, p. 285-298, 2021.
A002	HENNAH, N. A novel practical pedagogy for terminal assessment. Chemistry Education Research and Practice , v. 20, n. 1, p. 95-106, 2019.
A003	AYDIN, S. A science faculty's transformation of nature of science understanding into his teaching graduate level chemistry course. Chemistry Education Research and Practice , v. 16, n. 1, p. 133-142, 2015.
A004	SCOTT, F. J. A simulated peer-assessment approach to improving student performance in chemical calculations. Chemistry Education Research and Practice , v. 15, n. 4, p. 568-575, 2014.
A005	NASSON, I.; DORI, Y. J. A three-attribute transfer skills framework—part II: applying and assessing the model in science education. Chemistry Education Research and Practice , v. 16, n. 1, p. 154-167, 2015.
A006	TAN, K. C. D. <i>et al.</i> A web-based ionisation energy diagnostic instrument: exploiting the affordances of technology. Chemistry Education Research and Practice , v. 20, n. 2, p. 412-427, 2019.
A007	ROCABADO, G. A. <i>et al.</i> Addressing diversity and inclusion through group comparisons: a primer on measurement invariance testing. Chemistry Education Research and Practice , v. 21, n. 3, p. 969-988, 2020.
A008	COX Jr, C. T. <i>et al.</i> Advancements in curriculum and assessment by the use of IMMEX technology in the organic laboratory. Chemistry Education Research and Practice , v. 9, n. 2, p. 163-168, 2008.
A009	SANTOS-DÍAZ, S.; TOWNS, M. H. An all-female graduate student organization participating in chemistry outreach: a case study characterizing leadership in the community of practice. Chemistry Education Research and Practice , v. 22, n. 2, p. 532-553, 2021.
A010	SENDUR, G. An examination of pre-service chemistry teachers' meaningful understanding and learning difficulties about aromatic compounds using a systemic assessment questions diagram. Chemistry Education Research and Practice , v. 21, n. 1, p. 113-140, 2020.
A011	KIM, T.; WRIGHT, L. K.; MILLER, K. An examination of students' perceptions of the Kekulé resonance representation using a perceptual learning theory lens. Chemistry Education Research and Practice , v. 20, n. 4, p. 659-666, 2019.

A012	RALPH, V. R.; LEWIS, S. E. An explanative basis for the differential performance of students with low math aptitude in general chemistry. Chemistry Education Research and Practice , v. 20, n. 3, p. 570-593, 2019.
A013	JOHNSTON, P. R.; WATTERS, D. J.; BROWN, C. L.; LOUGHLIN, W. A. An investigation into student perceptions towards mathematics and their performance in first year chemistry: introduction of online maths skills support. Chemistry Education Research and Practice , v. 17, n. 4, p. 1203-1214, 2016.
A014	SHAH, L.; RODRIGUEZ, C. A.; BARTOLI, M.; RUSHTON, G. T. Analysing the impact of a discussion-oriented curriculum on first-year general chemistry students' conceptions of relative acidity. Chemistry Education Research and Practice , v. 19, n. 2, p. 543-557, 2018.
A015	RUSMINI, R.; SUYONO, S.; AGUSTINI, R. Analysis of science process skills of chemical education students through self project based learning (SjBL) in the pandemic COVID 19 era. JOTSE , v. 11, n. 2, p. 371-387, 2021.
A016	SCHMIDT-MCCORMACK, J. A. <i>et al.</i> Analysis of the role of a writing-to-learn assignment in student understanding of organic acid–base concepts. Chemistry Education Research and Practice , v. 20, n. 2, p. 383-398, 2019.
A017	WANG, Y.; LEWIS, S. E. Analytical chemistry students' explanatory statements in the context of their corresponding lecture. Chemistry Education Research and Practice , v. 21, n. 4, p. 1183-1198, 2020.
A018	DOOD, A. J.; DOOD, J. C.; DE ARELLANO, D. C. R.; FIELDS, K. B.; RAKER, J. R. Analyzing explanations of substitution reactions using lexical analysis and logistic regression techniques. Chemistry Education Research and Practice , v. 21, n. 1, p. 267-286, 2020.
A019	MOON, A.; MOELLER, R.; GERE, A. R.; SHULTZ, G. V. Application and testing of a framework for characterizing the quality of scientific reasoning in chemistry students' writing on ocean acidification. Chemistry Education Research and Practice , v. 20, n. 3, p. 484-494, 2019.
A020	COOPER, A. K.; OLIVER-HOYO, M. T. Argument construction in understanding noncovalent interactions: a comparison of two argumentation frameworks. Chemistry Education Research and Practice , v. 17, n. 4, p. 1006-1018, 2016.
A021	ROOTMAN-LE GRANGE, I.; BLACKIE, M. A. Assessing assessment: in pursuit of meaningful learning. Chemistry Education Research and Practice , v. 19, n. 2, p. 484-490, 2018.
A022	YAMAN, F.; AYAS, A. Assessing changes in high school students' conceptual understanding through concept maps before and after the computer-based predict–observe–explain (CB-POE) tasks on acid–base chemistry at the secondary level. Chemistry Education Research and Practice , v. 16, n. 4, p. 843-855, 2015.
A023	HOLME, T.; MURPHY, K. Assessing conceptual and algorithmic knowledge in general chemistry with ACS exams. Journal of Chemical Education , v. 88, n. 9, p. 1217-1222, 2011.

A024	TAL, M.; HERSCOVITZ, O.; DORI, Y. J. Assessing teachers' knowledge: incorporating context-based learning in chemistry. Chemistry Education Research and Practice , v. 22, n. 4, p. 1003-1019, 2021.
A025	BERNHOLT, S.; PARCHMANN, I. Assessing the complexity of students' knowledge in chemistry. Chemistry Education Research and Practice , v. 12, n. 2, p. 167-173, 2011.
A026	GIBBONS, R. E.; REED, J. J.; SRINIVASAN, S.; VILLAFANE, S. M.; LAGA, E.; VEGA, J.; <i>et al.</i> Assessment in postsecondary chemistry education: a comparison of course types. Assessment Update , v. 30, n. 3, p. 1-16, 2018.
A027	JAGODZIŃSKI, P.; WOLSKI, R. Assessment of application technology of natural user interfaces in the creation of a virtual chemical laboratory. Journal of Science Education and Technology , v. 24, n. 1, p. 16-28, 2015.
A028	MORAIS, C.; ARAÚJO, J. L. Awakening to chemistry through storytelling and practical activities: middle school students interacting with pre-school children. Chemistry Education Research and Practice , v. 20, n. 1, p. 302-315, 2019.
A029	CHEE, Y. S.; TAN, K. C. D. Becoming chemists through game-based inquiry learning: The case of legends of Alkhimia. 2012.
A030	ATIEH, E. L.; YORK, D. M.; MUÑIZ, M. N. Beneath the surface: an investigation of general chemistry students' study skills to predict course outcomes. Journal of Chemical Education , v. 98, n. 2, p. 281-292, 2020.
A031	KOUTALIDI, S.; SCOULLOS, M. Biogeochemical cycles for combining chemical knowledge and ESD issues in Greek secondary schools part I: designing the didactic materials. Chemistry Education Research and Practice , v. 17, n. 1, p. 10-23, 2016.
A032	REES, S. W.; KIND, V.; NEWTON, D. Can language focussed activities improve understanding of chemical language in non-traditional students? Chemistry Education Research and Practice , v. 19, n. 3, p. 755-766, 2018.
A033	WEINRICH, M. L.; SEVIAN, H. Capturing students' abstraction while solving organic reaction mechanism problems across a semester. Chemistry Education Research and Practice , v. 18, n. 1, p. 169-190, 2017.
A034	TASHIRO, J.; PARGA, D.; POLLARD, J.; TALANQUER, V. Characterizing change in students' self-assessments of understanding when engaged in instructional activities. Chemistry Education Research and Practice , v. 22, n. 3, p. 662-682, 2021.
A035	HARSHMAN, J.; YEZIERSKI, E. Characterizing high school chemistry teachers' use of assessment data via latent class analysis. Chemistry Education Research and Practice , v. 17, n. 2, p. 296-308, 2016.
A036	SAITTA, E. K. H.; BOWDON, M. A.; GEIGER, C. L. Incorporating service-learning, technology, and research supportive teaching techniques into the university chemistry classroom. Journal of Science Education and Technology , v. 20, n. 6, p. 790-795, 2011.
A037	SCHAFER, A. G.; YEZIERSKI, E. J. Chemistry critical friendships: investigating chemistry-specific discourse within a domain-general

	discussion of best practices for inquiry assessments. Chemistry Education Research and Practice , v. 21, n. 1, p. 452-468, 2020.
A038	POPOVA, M.; JONES, T. Chemistry instructors' intentions toward developing, teaching, and assessing student representational competence skills. Chemistry Education Research and Practice , v. 22, n. 3, p. 733-748, 2021.
A039	VOGELZANG, J.; ADMIRAAL, W. F. Classroom action research on formative assessment in a context-based chemistry course. Educational Action Research , v. 25, n. 1, p. 155-166, 2017.
A040	DEMIRDÖĞEN, B.; UZUNTIRYAKI-KONDAKÇI, E. Closing the gap between beliefs and practice: Change of pre-service chemistry teachers' orientations during a PCK-based NOS course. Chemistry Education Research and Practice , v. 17, n. 4, p. 818-841, 2016.
A041	DING, N.; HARSKAMP, E. G. Collaboration and peer tutoring in chemistry laboratory education. International Journal of Science Education , v. 33, n. 6, p. 839-863, 2011.
A042	YOUNG, K. R.; SCHAFFER, H. E.; JAMES, J. B.; GALLARDO-WILLIAMS, M. T. Tired of failing students? Improving student learning using detailed and automated individualized feedback in a large introductory science course. Innovative Higher Education , v. 46, n. 2, p. 133-151, 2021.
A043	NYACHWAYA, J. M.; WARFA, A. R. M.; ROEHRIG, G. H.; SCHNEIDER, J. L. College chemistry students' use of memorized algorithms in chemical reactions. Chemistry Education Research and Practice , v. 15, n. 1, p. 81-93, 2014.
A044	ELMGREN, M.; HO, F.; ÅKESSON, E.; SCHMID, S.; TOWNS, M. Comparison and evaluation of learning outcomes from an international perspective: Development of a best-practice process. Journal of Chemical Education , v. 92, n. 3, p. 427-432, 2015.
A045	BUSTOS-WORKS, C.; WHILES LILLIG, J.; CLARK, C.; DAUBENMIRE, P.; CLAESGENS, J.; SHUSTERMAN, A.; <i>et al.</i> Moving toward inclusivity in chemistry by developing data-based instructional tasks aimed at increasing students' self-perception as capable learners who belong in STEM. Journal of Chemical Education , v. 99, n. 1, p. 177-184, 2021.
A046	CANELAS, D. A.; HILL, J. L.; NOVICKI, A. Cooperative learning in organic chemistry increases student assessment of learning gains in key transferable skills. Chemistry Education Research and Practice , v. 18, n. 3, p. 441-456, 2017.
A047	RODRIGUEZ, J. M. G.; BAIN, K.; TOWNS, M. H.; ELMGREN, M.; HO, F. M. Covariational reasoning and mathematical narratives: Investigating students' understanding of graphs in chemical kinetics. Chemistry Education Research and Practice , v. 20, n. 1, p. 107-119, 2019.
A048	WARFA, A. R. M.; ODOWA, N. Creative exercises (CEs) in the biochemistry domain: an analysis of students' linking of chemical and biochemical concepts. Chemistry Education Research and Practice , v. 16, n. 4, p. 747-757, 2015.
A049	ALI, N.; ULLAH, S.; RAEES, M. Interactive cube for effective demonstration of virtual periodic table. Education and Information Technologies , v. 27, n. 2, p. 1635-1654, 2022.

A050	EL-KHAWAS, E. Defining the role of academics in accountability. Higher Education Management and Policy , v. 21, n. 1, p. 1-12, 2009.
A051	ANTONOGLOU, L. D.; CHARISTOS, N. D.; SIGALAS, M. P. Design, development and implementation of a technology enhanced hybrid course on molecular symmetry: Students' outcomes and attitudes. Chemistry Education Research and Practice , v. 12, n. 4, p. 454-468, 2011.
A052	YUSUF, M.; TAYLOR, P. C.; DAMANHURI, M. I. M. Designing critical pedagogy to counteract the hegemonic culture of the traditional chemistry classroom. Issues in Educational Research , v. 27, n. 1, p. 168-184, 2017.
A053	HANSON, R. Designing E-Content for Teaching Basic Analytical Chemistry in Higher Education: A Baseline Study. Science Education International , v. 31, n. 1, p. 22-28, 2020.
A054	HARSH, J. A. Designing performance-based measures to assess the scientific thinking skills of chemistry undergraduate researchers. Chemistry Education Research and Practice , v. 17, n. 4, p. 808-817, 2016.
A055	ARMSTRONG, L. B.; RIVAS, M. C.; ZHOU, Z.; IRIE, L. M.; KERSTIENS, G. A.; ROBAK, M. T.; <i>et al.</i> Developing a Green Chemistry Focused General Chemistry Laboratory Curriculum: What Do Students Understand and Value about Green Chemistry? Journal of Chemical Education , v. 96, n. 11, p. 2410-2419, 2019.
A056	SEERY, M. K.; AGUSTIAN, H. Y.; DOIDGE, E. D.; KUCHARSKI, M. M.; O'CONNOR, H. M.; PRICE, A. Developing laboratory skills by incorporating peer-review and digital badges. Chemistry Education Research and Practice , v. 18, n. 3, p. 403-419, 2017.
A057	LOERTSCHER, J.; LEWIS, J. E.; MERCER, A. M.; MINDERHOUT, V. Development and use of a construct map framework to support teaching and assessment of noncovalent interactions in a biochemical context. Chemistry Education Research and Practice , v. 19, n. 4, p. 1151-1165, 2018.
A058	CHRISTIAN, B. N.; YEZIERSKI, E. J. Development and validation of an instrument to measure student knowledge gains for chemical and physical change for grades 6–8. Chemistry Education Research and Practice , v. 13, n. 3, p. 384-393, 2012.
A059	DANCZAK, S. M.; THOMPSON, C. D.; OVERTON, T. L. Development and validation of an instrument to measure undergraduate chemistry students' critical thinking skills. Chemistry Education Research and Practice , v. 21, n. 1, p. 62-78, 2020.
A060	YIK, B. J.; DOOD, A. J.; DE ARELLANO, D. C. R.; FIELDS, K. B.; RAKER, J. R. Development of a machine learning-based tool to evaluate correct Lewis acid–base model use in written responses to open-ended formative assessment items. Chemistry Education Research and Practice , v. 22, n. 4, p. 866-885, 2021.
A061	LU, H.; JIANG, Y.; BI, H. Development of a measurement instrument to assess students' proficiency levels regarding galvanic cells. Chemistry Education Research and Practice , v. 21, n. 2, p. 655-667, 2020.
A062	EKIZ-KIRAN, B.; BOZ, Y.; OZTAY, E. S. Development of pre-service teachers' pedagogical content knowledge through a PCK-based school

	experience course. Chemistry Education Research and Practice , v. 22, n. 2, p. 415-430, 2021.
A063	BRANDRIET, A. R.; XU, X.; BRETZ, S. L.; LEWIS, J. E. Diagnosing changes in attitude in first-year college chemistry students with a shortened version of Bauer's semantic differential. Chemistry Education Research and Practice , v. 12, n. 2, p. 271-278, 2011.
A064	HRIN, T.; MILENKOVIĆ, D.; SEGEDINAC, M. Diagnosing the quality of high school students' and pre-service chemistry teachers' cognitive structures in organic chemistry by using students' generated systemic synthesis questions. Chemistry Education Research and Practice , v. 19, n. 1, p. 305-318, 2018.
A065	ÖZALP, D.; KAHVECI, A. Diagnostic assessment of student misconceptions about the particulate nature of matter from ontological perspective. Chemistry Education Research and Practice , v. 16, n. 3, p. 619-639, 2015.
A066	ZHANG, L.; WANG, L.; TREAGUST, D. F. Discipline-specific cognitive factors that influence grade 9 students' performance in chemistry. Chemistry Education Research and Practice , v. 22, n. 4, p. 813-841, 2021.
A067	GILEWSKI, A.; MALLORY, E.; SANDOVAL, M.; LITVAK, M.; YE, L. Does linking help? Effects and student perceptions of a learner-centered assessment implemented in introductory chemistry. Chemistry Education Research and Practice , v. 20, n. 2, p. 399-411, 2019.
A068	DALGARNO, B.; BISHOP, A. G.; ADLONG, W.; BEDGOOD Jr, D. R. Effectiveness of a virtual laboratory as a preparatory resource for distance education chemistry students. Computers & Education , v. 53, n. 3, p. 853-865, 2009.
A069	HOUSEKNECHT, J. B.; BACHINSKI, G. J.; MILLER, M. H.; WHITE, S. A.; ANDREWS, D. M. Effectiveness of the active learning in organic chemistry faculty development workshops. Chemistry Education Research and Practice , v. 21, n. 1, p. 387-398, 2020.
A070	PETTERSON, M. N.; WATTS, F. M.; SNYDER-WHITE, E. P.; ARCHER, S. R.; SHULTZ, G. V.; FINKENSTAEDT-QUINN, S. A. Eliciting student thinking about acid-base reactions via app and paper-pencil based problem solving. Chemistry Education Research and Practice , v. 21, n. 3, p. 878-892, 2020.
A071	HANSON, R. Employing Microscience Equipment to Promote Chemistry Education Through Constructivist Hands-and Minds-on Activities. Interchange , v. 52, n. 4, p. 521-544, 2021.
A072	STULL, A. T.; GAINER, M. J.; HEGARTY, M. Learning by enacting: The role of embodiment in chemistry education. Learning and Instruction , v. 55, p. 80-92, 2018.
A073	LAU, P. N. Enhancing formative and self-assessment with video playback to improve critique skills in a titration laboratory. Chemistry Education Research and Practice , v. 21, n. 1, p. 178-188, 2020.
A074	GHANI, I. A.; IBRAHIM, N. H.; YAHAYA, N. A.; SURIF, J. Enhancing students' HOTS in laboratory educational activity by using concept map as an alternative assessment tool. Chemistry Education Research and Practice , v. 18, n. 4, p. 849-874, 2017.

A075	HOLME, T.; BRETZ, S. L.; COOPER, M.; LEWIS, J.; PAEK, P.; PIENTA, N.; <i>et al.</i> Enhancing the role of assessment in curriculum reform in chemistry. Chemistry Education Research and Practice , v. 11, n. 2, p. 92-97, 2010.
A076	CARLE, M. S.; FLYNN, A. B. Essential learning outcomes for delocalization (resonance) concepts: How are they taught, practiced, and assessed in organic chemistry? Chemistry Education Research and Practice , v. 21, n. 2, p. 622-637, 2020.
A077	LEWIS, S. E.; SHAW, J. L.; FREEMAN, K. A. Establishing open-ended assessments: investigating the validity of creative exercises. Chemistry Education Research and Practice , v. 12, n. 2, p. 158-166, 2011.
A078	NEILES, K. Y.; TODD, I.; BUNCE, D. M. Establishing the validity of using network analysis software for measuring students' mental storage of chemistry concepts. Journal of Chemical Education , v. 93, n. 5, p. 821-831, 2016.
A079	ÇİÇEK, Ö.; ILHAN, N. A. İ. L. Evaluating interest in acids–bases: development of an acid–base interest scale (ABIS) and assessment of pre-service science teachers' interest. Chemistry Education Research and Practice , v. 18, n. 4, p. 630-640, 2017.
A080	HARSH, J.; ESTEB, J. J.; MALTese, A. V. Evaluating the development of chemistry undergraduate researchers' scientific thinking skills using performance-data: first findings from the performance assessment of undergraduate research (PURE) instrument. Chemistry Education Research and Practice , v. 18, n. 3, p. 472-485, 2017.
A081	YÜKSEL, M. Evaluating the Effectiveness of the Chemistry Education by Using the Analytic Hierarchy Process. International Education Studies , v. 5, n. 5, p. 79-91, 2012.
A082	EL-FARARGY, N. Evaluation of a chemistry curriculum intervention using the Perry model of intellectual development. Chemistry Education Research and Practice , v. 11, n. 2, p. 98-106, 2010.
A083	SCHULTZ, M. <i>et al.</i> Evaluation of diagnostic tools that tertiary teachers can apply to profile their students' conceptions. International Journal of Science Education , v. 39, n. 5, p. 565-586, 2017.
A084	AYDIN, S.; FRIEDRICHSEN, P. M.; BOZ, Y.; HANUSCIN, D. L. Examination of the topic-specific nature of pedagogical content knowledge in teaching electrochemical cells and nuclear reactions. Chemistry Education Research and Practice , v. 15, n. 4, p. 658-674, 2014.
A085	CHEN, B.; WEI, B. Examining chemistry teachers' use of curriculum materials: in view of teachers' pedagogical content knowledge. Chemistry Education Research and Practice , v. 16, n. 2, p. 260-272, 2015.
A086	LEWIS, S. E. Examining evidence for external and consequential validity of the first term general chemistry exam from the ACS Examinations Institute. Journal of Chemical Education , v. 91, n. 6, p. 793-799, 2014.
A087	SUPALO, C. A. <i>et al.</i> Examining the use of adaptive technologies to increase the hands-on participation of students with blindness or low vision in secondary-school chemistry and physics. Chemistry Education Research and Practice , v. 17, n. 4, p. 1174-1189, 2016.

A088	ANWAR, Y. A. S.; MUTI'AH, M. A. Exploration of the scientific papers and self-assessment of students using the COVID-19 case on biochemistry course. Biochemistry and Molecular Biology Education , v. 49, n. 3, p. 326-332, 2021.
A089	VILLAFANE, S. M.; LEWIS, J. E. Exploring a measure of science attitude for different groups of students enrolled in introductory college chemistry. Chemistry Education Research and Practice , v. 17, n. 4, p. 731-742, 2016.
A090	RYOO, K.; TOUTKOUSHIAN, E.; BEDELL, K. Exploring different types of assessment items to measure linguistically diverse students' understanding of energy and matter in chemistry. Chemistry Education Research and Practice , v. 19, n. 1, p. 149-166, 2018.
A091	ZANE, M.; TUCCI, V. K. Exploring the information literacy needs and values of high school chemistry teachers. Journal of Chemical Education , v. 93, n. 3, p. 406-412, 2016.
A092	SIWEYA, H. J.; LETSOALO, P. Formative assessment by first-year chemistry students as predictor of success in summative assessment at a South African university. Chemistry Education Research and Practice , v. 15, n. 4, p. 541-549, 2014.
A093	LAZENBY, K. et al. From Ideas to Items: A Primer on the Development of Ordered Multiple-Choice Items for Investigating the Progression of Learning in Higher Education STEM. Journal of Chemical Education , v. 98, n. 3, p. 714-729, 2021.
A094	DONNELLY, J.; HERNÁNDEZ, F. E. Fusing a reversed and informal learning scheme and space: student perceptions of active learning in physical chemistry. Chemistry Education Research and Practice , v. 19, n. 2, p. 520-532, 2018.
A095	MA, J. Getting a Bigger Picture in Less Time: Viewing Curriculum Reform in a Chinese Graduate Chemistry Program through the Lens of an Organic Structure Analysis Course. Journal of Chemical Education , v. 88, n. 12, p. 1639-1643, 2011.
A096	HARSHMAN, J.; YEZIERSKI, E. Guiding teaching with assessments: high school chemistry teachers' use of data-driven inquiry. Chemistry Education Research and Practice , v. 16, n. 1, p. 93-103, 2015.
A097	KINGIR, S.; GEBAN, O.; GUNEL, M. How does the science writing heuristic approach affect students' performances of different academic achievement levels? A case for high school chemistry. Chemistry Education Research and Practice , v. 13, n. 4, p. 428-436, 2012.
A098	RAVI, M.; PUENTE-URBINA, A.; VAN BOKHOVEN, J. A. Identifying Opportunities to Promote Systems Thinking in Catalysis Education. Journal of Chemical Education , v. 98, n. 5, p. 1583-1593, 2021.
A099	CHEN, Y. C.; WILSON, K.; LIN, H. S. Identifying the challenging characteristics of systems thinking encountered by undergraduate students in chemistry problem-solving of gas laws. Chemistry Education Research and Practice , v. 20, n. 3, p. 594-605, 2019.
A100	NOTTIS, K. E.; KASTNER, M. E. The effect of instructional modality and prior knowledge on learning point group symmetry. Journal of Science Education and Technology , v. 14, n. 1, p. 51-58, 2005.

A101	WILLIAMSON, V. M. et al. Impact of basic arithmetic skills on success in first-semester general chemistry. Chemistry Education Research and Practice , v. 21, n. 1, p. 51-61, 2020.
A102	APUGLIESE, A.; LEWIS, S. E. Impact of instructional decisions on the effectiveness of cooperative learning in chemistry through meta-analysis. Chemistry Education Research and Practice , v. 18, n. 1, p. 271-278, 2017.
A103	BEHMKE, D. A.; ATWOOD, C. H. Implementation and assessment of Cognitive Load Theory (CLT) based questions in an electronic homework and testing system. Chemistry Education Research and Practice , v. 14, n. 3, p. 247-256, 2013.
A104	CHILDS, P. E. Improving chemical education: turning research into effective practice. Chemistry Education Research and Practice , v. 10, n. 3, p. 189-203, 2009.
A105	WILLIAMS, N. A.; BLAND, W.; CHRISTIE, G. Improving student achievement and satisfaction by adopting a blended learning approach to inorganic chemistry. Chemistry Education Research and Practice , v. 9, n. 1, p. 43-50, 2008.
A106	HORN, S.; HERNICK, M. Improving student understanding of lipids concepts in a biochemistry course using test-enhanced learning. Chemistry Education Research and Practice , v. 16, n. 4, p. 918-928, 2015.
A107	WATSON, S. W.; DUBROVSKIY, A. V.; PETERS, M. L. Increasing chemistry students' knowledge, confidence, and conceptual understanding of pH using a collaborative computer pH simulation. Chemistry Education Research and Practice , v. 21, n. 2, p. 528-535, 2020.
A108	PANTIWATI, Y.; WAHYUNI, S.; PERMANA, F. H. Instructional model of natural science in junior high schools, Batu-Malang. Journal of Education and Practice , v. 8, n. 8, 2017.
A109	TAAGEPERA, M. et al. Integrating symmetry in stereochemical analysis in introductory organic chemistry. Chemistry Education Research and Practice , v. 12, n. 3, p. 322-330, 2011.
A110	GRAULICH, N. Intuitive judgments govern students' answering patterns in multiple-choice exercises in organic chemistry. Journal of Chemical Education , v. 92, n. 2, p. 205-211, 2015.
A111	SCHAFER, A. G.; YEZIERSKI, E. J. Investigating high school chemistry teachers' assessment item generation processes for a solubility lab. Chemistry Education Research and Practice , v. 22, n. 1, p. 214-225, 2021.
A112	SANABRIA-RÍOS, D.; BRETZ, S. L. Investigating the relationship between faculty cognitive expectations about learning chemistry and the construction of exam questions. Chemistry Education Research and Practice , v. 11, n. 3, p. 212-217, 2010.
A113	PULLEN, R.; THICKETT, S. C.; BISSEMBER, A. C. Investigating the viability of a competency-based, qualitative laboratory assessment model in first-year undergraduate chemistry. Chemistry Education Research and Practice , v. 19, n. 2, p. 629-637, 2018.
A114	FLYNN, A. B.; FEATHERSTONE, R. B. Language of mechanisms: exam analysis reveals students' strengths, strategies, and errors when using

	the electron-pushing formalism (curved arrows) in new reactions. Chemistry Education Research and Practice , v. 18, n. 1, p. 64-77, 2017.
A115	VISHNUMOLAKALA, V. R. et al. Latent constructs of the students' assessment of their learning gains instrument following instruction in stereochemistry. Chemistry Education Research and Practice , v. 17, n. 2, p. 309-319, 2016.
A116	DANGUR, V. et al. Learning quantum chemistry via a visual-conceptual approach: students' bidirectional textual and visual understanding. Chemistry Education Research and Practice , v. 15, n. 3, p. 297-310, 2014.
A117	RYAN, B. J. Line up, line up: using technology to align and enhance peer learning and assessment in a student centred foundation organic chemistry module. Chemistry Education Research and Practice , v. 14, n. 3, p. 229-238, 2013.
A118	YE, L.; LEWIS, S. E. Looking for links: examining student responses in creative exercises for evidence of linking chemistry concepts. Chemistry Education Research and Practice , v. 15, n. 4, p. 576-586, 2014.
A119	KOUSA, P.; KAVONIUS, R.; AKSELA, M. Low-achieving students' attitudes towards learning chemistry and chemistry teaching methods. Chemistry Education Research and Practice , v. 19, n. 2, p. 431-441, 2018.
A120	KISHBAUGH, T. L. et al. Measuring beyond content: a rubric bank for assessing skills in authentic research assignments in the sciences. Chemistry Education Research and Practice , v. 13, n. 3, p. 268-276, 2012.
A121	MATHABATHE, K. C.; POTGIETER, M. Metacognitive monitoring and learning gain in foundation chemistry. Chemistry Education Research and Practice , v. 15, n. 1, p. 94-104, 2014.
A122	MACK, M. R.; HENSEN, C.; BARBERA, J. Metrics and methods used to compare student performance data in chemistry education research articles. Journal of Chemical Education , v. 96, n. 3, p. 401-413, 2019.
A123	MARTIN, S. N.; SCANTLEBURY, K. More than a conversation: Using cogenerative dialogues in the professional development of high school chemistry teachers. Educational Assessment, Evaluation and Accountability , v. 21, n. 2, p. 119-136, 2009.
A124	TURNER, M. et al. Motivations for and barriers to the implementation of diagnostic assessment practices—a case study. Chemistry Education Research and Practice , v. 12, n. 2, p. 142-157, 2011.
A125	FOX, L. J.; ROEHRIG, G. H. Nationwide survey of the undergraduate physical chemistry course. Journal of Chemical Education , v. 92, n. 9, p. 1456-1465, 2015.
A126	KRAUSE, M. et al. On the development and assessment of a computer-based learning and assessment environment for the transition from lower to upper secondary chemistry education. Chemistry Education Research and Practice , v. 14, n. 3, p. 345-353, 2013.
A127	O'DWYER, A.; CHILDS, P. Organic chemistry in action! Developing an intervention program for introductory organic chemistry to improve learners' understanding, interest, and attitudes. Journal of Chemical Education , v. 91, n. 7, p. 987-993, 2014.

A128	CARROLL, M. K. Moving from Recommendations to Innovations: Increasing the Relevancy and Effectiveness of Chemistry Education. Journal of Chemical Education , v. 90, n. 7, p. 816-819, 2013.
A129	GALLOWAY, K. R.; LEUNG, M. W.; FLYNN, A. B. Patterns of reactions: a card sort task to investigate students' organization of organic chemistry reactions. Chemistry Education Research and Practice , v. 20, n. 1, p. 30-52, 2019.
A130	GOK, T.; GOK, O. Peer instruction in chemistry education: Assessment of students' learning strategies, conceptual learning and problem solving. In: Asia-Pacific Forum on Science Learning & Teaching , v. 17, n. 1, jun. 2016.
A131	LAWRIE, G. A. et al. Personal journeys of teachers: an investigation of the development of teacher professional knowledge and skill by expert tertiary chemistry teachers. Chemistry Education Research and Practice , v. 20, n. 1, p. 132-145, 2019.
A132	NIELSEN, W.; HOBAN, G.; HYLAND, C. J. Pharmacology students' perceptions of creating multimodal digital explanations. Chemistry Education Research and Practice , v. 18, n. 2, p. 329-339, 2017.
A133	ALLENBAUGH, R. J.; HERRERA, K. M. Pre-assessment and peer tutoring as measures to improve performance in gateway general chemistry classes. Chemistry Education Research and Practice , v. 15, n. 4, p. 620-627, 2014.
A134	KARARO, A. T. et al. Predictions and constructing explanations: an investigation into introductory chemistry students' understanding of structure–property relationships. Chemistry Education Research and Practice , v. 20, n. 1, p. 316-328, 2019.
A135	BEKTAS, O. et al. Pre-service chemistry teachers' pedagogical content knowledge of the nature of science in the particle nature of matter. Chemistry Education Research and Practice , v. 14, n. 2, p. 201-213, 2013.
A136	GÜLTEPE, N. Pre-service Chemistry Teachers' Understanding About Equilibria in Acid-Base Solutions. Pedagogical Research , v. 6, n. 4, 2021.
A137	ZEMEL, Y.; SHWARTZ, G.; AVARGIL, S. Preservice teachers' enactment of formative assessment using rubrics in the inquiry-based chemistry laboratory. Chemistry Education Research and Practice , v. 22, n. 4, p. 1074-1092, 2021.
A138	CRUJEIRAS-PÉREZ, B.; BROCOS, P. Pre-service teachers' use of epistemic criteria in the assessment of scientific procedures for identifying microplastics in beach sand. Chemistry Education Research and Practice , v. 22, n. 2, p. 237-246, 2021.
A139	UNDERSANDER, M. A. et al. Probing the question order effect while developing a chemistry concept inventory. Chemistry Education Research and Practice , v. 18, n. 1, p. 45-54, 2017.
A140	OLIVER-HOYO, M.; BABILONIA-ROSA, M. A. Promotion of spatial skills in chemistry and biochemistry education at the college level. Journal of Chemical Education , v. 94, n. 8, p. 996-1006, 2017.
A141	WREN, D.; BARBERA, J. Psychometric analysis of the thermochemistry concept inventory. Chemistry Education Research and Practice , v. 15, n. 3, p. 380-390, 2014.

A142	DENG, J. M.; FLYNN, A. B. Reasoning, granularity, and comparisons in students' arguments on two organic chemistry items. Chemistry Education Research and Practice , v. 22, n. 3, p. 749-771, 2021.
A143	CHIANG, K. H. Relationship between research and teaching in doctoral education in UK universities. Higher Education Policy , v. 17, n. 1, p. 71-88, 2004.
A144	COOPER, M. M.; SANDI-URENA, S.; STEVENS, R. Reliable multi method assessment of metacognition use in chemistry problem solving. Chemistry Education Research and Practice , v. 9, n. 1, p. 18-24, 2008.
A145	DENG, Y.; WANG, H. Research on evaluation of Chinese students' competence in written scientific argumentation in the context of chemistry. Chemistry Education Research and Practice , v. 18, n. 1, p. 127-150, 2017.
A146	EMENIKE, M. E. et al. Results from a national needs assessment survey: A view of assessment efforts within chemistry departments. Journal of Chemical Education , v. 90, n. 5, p. 561-567, 2013.
A147	LEVY, S. T.; WILENSKY, U. Students' learning with the connected chemistry (CC1) curriculum: Navigating the complexities of the particulate world. Journal of Science Education and Technology , v. 18, n. 3, p. 243-254, 2009.
A148	TALBERT, L. E. et al. Revisiting the use of concept maps in a large enrollment general chemistry course: implementation and assessment. Chemistry Education Research and Practice , v. 21, n. 1, p. 37-50, 2020.
A149	PRINS, G. T.; BULTE, A. M.; VAN DRIEL, J. H.; PILOT, A. Selection of authentic modelling practices as contexts for chemistry education. International Journal of Science Education , v. 30, n. 14, p. 1867-1890, 2008.
A150	EATON, A. C.; DELANEY, S.; SCHULTZ, M. Situating sustainable development within secondary chemistry education via systems thinking: a depth study approach. Journal of Chemical Education , v. 96, n. 12, p. 2968-2974, 2019.
A151	RAJE, S.; STITZEL, S. Strategies for effective assessments while ensuring academic integrity in general chemistry courses during COVID-19. Journal of Chemical Education , v. 97, n. 9, p. 3436-3440, 2020.
A152	FLYNN, A. B. Structure and evaluation of flipped chemistry courses: organic & spectroscopy, large and small, first to third year, English and French. Chemistry Education Research and Practice , v. 16, n. 2, p. 198-211, 2015.
A153	SCHNEIDER, J. L.; RUDER, S. M.; BAUER, C. F. Student perceptions of immediate feedback testing in student centered chemistry classes. Chemistry Education Research and Practice , v. 19, n. 2, p. 442-451, 2018.
A154	BROMAN, K.; PARCHMANN, I. Students' application of chemical concepts when solving chemistry problems in different contexts. Chemistry Education Research and Practice , v. 15, n. 4, p. 516-529, 2014.
A155	WALPUSKI, M.; ROPHOL, M.; SUMFLETH, E. Students' knowledge about chemical reactions–development and analysis of standard-based

	test items. Chemistry Education Research and Practice , v. 12, n. 2, p. 174-183, 2011.
A156	HUANG, J. Successes and challenges: online teaching and learning of chemistry in higher education in China in the time of COVID-19. Journal of Chemical Education , v. 97, n. 9, p. 2810-2814, 2020.
A157	WHEELER, L. B.; MAENG, J. L.; WHITWORTH, B. A. Teaching assistants' perceptions of a training to support an inquiry-based general chemistry laboratory course. Chemistry Education Research and Practice , v. 16, n. 4, p. 824-842, 2015.
A158	HARSHMAN, J.; YEZIERSKI, E. Test-retest reliability of the adaptive chemistry assessment survey for teachers: Measurement error and alternatives to correlation. Journal of Chemical Education , v. 93, n. 2, p. 239-247, 2016.
A159	SENOCAK, E.; BALOGLU, M. The adaptation and preliminary psychometric properties of the Derived Chemistry Anxiety Rating Scale. Chemistry Education Research and Practice , v. 15, n. 4, p. 800-806, 2014.
A160	NYACHWAYA, J. M. et al. The development of an open-ended drawing tool: an alternative diagnostic tool for assessing students' understanding of the particulate nature of matter. Chemistry Education Research and Practice , v. 12, n. 2, p. 121-132, 2011.
A161	SEGEDINAC, M. D. et al. The development of research in the field of chemistry education at the university of Novi Sad since the breakup of the Socialist Federal Republic of Yugoslavia. CEPS Journal , v. 10, n. 1, p. 103-124, 2020.
A162	PARK, M. et al. The effect of computer models as formative assessment on student understanding of the nature of models. Chemistry Education Research and Practice , v. 18, n. 4, p. 572-581, 2017.
A163	PAGE, M. F. et al. The effect of teaching the entire academic year of high school chemistry utilizing abstract reasoning. Chemistry Education Research and Practice , v. 19, n. 2, p. 500-507, 2018.
A164	FAUTCH, J. M. The flipped classroom for teaching organic chemistry in small classes: is it effective?. Chemistry Education Research and Practice , v. 16, n. 1, p. 179-186, 2015.
A165	YE, L. et al. The impact of coupling assessments on conceptual understanding and connection-making in chemical equilibrium and acid-base chemistry. Chemistry Education Research and Practice , v. 21, n. 3, p. 1000-1012, 2020.
A166	COLLINS, J. S.; OLESIK, S. V. The important role of chemistry department chairs and recommendations for actions they can enact to advance black student success. Journal of Chemical Education , v. 98, n. 7, p. 2209-2220, 2021.
A167	CHEN, K. et al. The individual experience of online chemistry teacher education in China: Coping with COVID-19 pandemic. Journal of Chemical Education , v. 97, n. 9, p. 3265-3270, 2020.
A168	ELMAS, R. et al. The intellectual demands of the intended chemistry curriculum in Czechia, Finland, and Turkey: a comparative analysis based on the revised Bloom's taxonomy. Chemistry Education Research and Practice , v. 21, n. 3, p. 839-851, 2020.

A169	ALMEIDA, P. A. et al. The interplay between students' perceptions of context and approaches to learning. Research Papers in Education , v. 26, n. 2, p. 149-169, 2011.
A170	AYDIN, S.; BOZ, Y. The nature of integration among PCK components: A case study of two experienced chemistry teachers. Chemistry Education Research and Practice , v. 14, n. 4, p. 615-624, 2013.
A171	PROPHET, R. B.; VLAARDINGERBROEK, B. The relevance of secondary school chemistry education in Botswana: a cognitive development status perspective. International Journal of Educational Development , v. 23, n. 3, p. 275-289, 2003.
A172	SOUTHAM, D. C. et al. The timing of an experiment in the laboratory program is crucial for the student laboratory experience: acylation of ferrocene as a case study. Chemistry Education Research and Practice , v. 14, n. 4, p. 476-484, 2013.
A173	GENCER, S.; AKKUS, H. The topic-specific nature of experienced chemistry teachers' pedagogical content knowledge in the topics of interactions between chemical species and states of matter. Chemistry Education Research and Practice , v. 22, n. 2, p. 498-512, 2021.
A174	KOEPEL, I. et al. Turning chemistry education on its head: Design, experience and evaluation of a learning-centred 'Modern Chemistry' subject. Journal of University Teaching & Learning Practice , v. 17, n. 3, p. 13, 2020.
A175	VILLAFANE, S. M. et al. Uncovering students' incorrect ideas about foundational concepts for biochemistry. Chemistry Education Research and Practice , v. 12, n. 2, p. 210-218, 2011.
A176	LAMICHHANE, R.; RECK, C.; MALTESE, A. V. Undergraduate chemistry students' misconceptions about reaction coordinate diagrams. Chemistry Education Research and Practice , v. 19, n. 3, p. 834-845, 2018.
A177	VLADUŠIĆ, R.; BUCAT, R.; OŽIĆ, M. Understanding of words and symbols by chemistry university students in Croatia. Chemistry Education Research and Practice , v. 17, n. 3, p. 474-488, 2016.
A178	IRBY, S. M. et al. Use of a card sort task to assess students' ability to coordinate three levels of representation in chemistry. Chemistry Education Research and Practice , v. 17, n. 2, p. 337-352, 2016.
A179	MURCIANO-CALLES, J. Use of kahoot for assessment in chemistry education: a comparative study. Journal of Chemical Education , v. 97, n. 11, p. 4209-4213, 2020.
A180	SEVIAN, H. et al. Use of representation mapping to capture abstraction in problem solving in different courses in chemistry. Chemistry Education Research and Practice , v. 16, n. 3, p. 429-446, 2015.
A181	DERMAN, A.; EILKS, I. Using a word association test for the assessment of high school students' cognitive structures on dissolution. Chemistry Education Research and Practice , v. 17, n. 4, p. 902-913, 2016.
A182	SMITH, K. C.; VILLARREAL, S. Using animations in identifying general chemistry students' misconceptions and evaluating their knowledge transfer relating to particle position in physical changes. Chemistry Education Research and Practice , v. 16, n. 2, p. 273-282, 2015.
A183	BURROWS, N. L.; MOORING, S. R. Using concept mapping to uncover students' knowledge structures of chemical bonding concepts.

	Chemistry Education Research and Practice , v. 16, n. 1, p. 53-66, 2015.
A184	SCOTT, D.; FIRTH, D. Using control charts early in the quantitative analysis laboratory curriculum. Journal of Chemical Education , v. 96, n. 5, p. 1037-1041, 2019.
A185	CHEUNG, D. Using diagnostic assessment to help teachers understand the chemistry of the lead-acid battery. Chemistry Education Research and Practice , v. 12, n. 2, p. 228-237, 2011.
A186	HERRMANN-ABELL, C. F.; DEBOER, G. E. Using distractor-driven standards-based multiple-choice assessments and Rasch modeling to investigate hierarchies of chemistry misconceptions and detect structural problems with individual items. Chemistry Education Research and Practice , v. 12, n. 2, p. 184-192, 2011.
A187	WALKER, J. P. et al. Using the laboratory to engage all students in science practices. Chemistry Education Research and Practice , v. 17, n. 4, p. 1098-1113, 2016.
A188	NAKIBOGLU, C. Using word associations for assessing non major science students' knowledge structure before and after general chemistry instruction: the case of atomic structure. Chemistry Education Research and Practice , v. 9, n. 4, p. 309-322, 2008.
A189	LOPEZ, E. et al. Validating the use of concept-mapping as a diagnostic assessment tool in organic chemistry: implications for teaching. Chemistry Education Research and Practice , v. 12, n. 2, p. 133-141, 2011.
A190	HERRIDGE, M.; TASHIRO, J.; TALANQUER, V. Variation in chemistry instructors' evaluations of student written responses and its impact on grading. Chemistry Education Research and Practice , v. 22, n. 4, p. 948-972, 2021.
A191	ERDMANN, M. A.; MARCH, J. L. Video reports as a novel alternate assessment in the undergraduate chemistry laboratory. Chemistry Education Research and Practice , v. 15, n. 4, p. 650-657, 2014.
A192	VENKATARAMAN, B. Visualization and interactivity in the teaching of chemistry to science and non-science students. Chemistry Education Research and Practice , v. 10, n. 1, p. 62-69, 2009.
A193	SCHAFER, A. G.; BORLAND, V. M.; YEZIERSKI, E. J. Visualizing chemistry teachers' enacted assessment design practices to better understand barriers to "best practices". Chemistry Education Research and Practice , v. 22, n. 2, p. 457-475, 2021.
A194	DAVENPORT, J. L.; RAFFERTY, A. N.; YARON, D. J. Whether and how authentic contexts using a virtual chemistry lab support learning. Journal of Chemical Education , v. 95, n. 8, p. 1250-1259, 2018.
A195	BALAKRISHNAN, S. Who Wants to Be a Chemist? Formative Assessment Made Fun and Engaging. Journal of Educational Technology , v. 15, n. 2, p. 1-7, 2018.
A196	STOWE, R. L. et al. You are what you assess: The case for emphasizing chemistry on chemistry assessments. Journal of Chemical Education , v. 98, n. 8, p. 2490-2495, 2021.
A197	ZOLLER, U. Algorithmic, LOCS and HOCS (chemistry) exam questions: Performance and attitudes of college students. International Journal of Science Education , v. 24, n. 2, p. 185-203, 2002.

A198	YORK, S. et al. Applications of systems thinking in STEM education. Journal of Chemical Education , v. 96, n. 12, p. 2742-2751, 2019.
A199	WILSON, M.; SCALISE, K. Assessment to improve learning in higher education: The BEAR Assessment System. Higher Education , v. 52, n. 4, p. 635-663, 2006.
A200	LOZANO, R.; WATSON, M. K. Chemistry education for sustainability: Assessing the chemistry curricula at Cardiff University. Educación Química , v. 24, n. 2, p. 184-192, 2013.
A201	SCHERER, R.; MEßINGER-KOPPELT, J.; TIEMANN, R. Developing a computer-based assessment of complex problem solving in Chemistry. International Journal of STEM Education , v. 1, n. 1, p. 1-15, 2014.
A202	VACHLIOTIS, T.; SALTA, K.; TZOUGRAKI, C. Developing Basic Systems Thinking Skills for Deeper Understanding of Chemistry Concepts in High School Students. Thinking Skills and Creativity , v. 41, p. 100881, 2021.
A203	ZOLLER, U. Education in environmental chemistry: Setting the agenda and recommending action. A workshop report summary. Journal of Chemical Education , v. 82, n. 8, p. 1237, 2005.
A204	BOWEN, R. S.; COOPER, M. M. Grading on a Curve as a Systemic Issue of Equity in Chemistry Education. Journal of Chemical Education , v. 99, n. 1, p. 185-194, 2021.
A205	NAGEL, M. L.; LINDSEY, B. A. Implementation of reasoning chain construction tasks to support student explanations in general chemistry. Journal of Chemical Education , v. 99, n. 2, p. 839-850, 2021.