

Uma revisão da literatura acerca da implementação da metodologia interativa de ensino *Peer Instruction* (1991 a 2015)

A literature review on the implementation of Peer Instruction interactive teaching method (1991 to 2015)

Maykon Gonçalves Müller*¹, Ives Solano Araujo², Eliane Angela Veit², Julie Schell³

¹Instituto Federal de Educação, Ciência e Tecnologia Sul-rio-grandense, Campus Visconde da Graça, Pelotas, RS, Brasil

²Instituto de Física, Universidade Federal do Rio Grande do Sul, Porto Alegre, RS, Brasil

³Universidade do Texas, Austin, EUA

Recebido em 17 de Janeiro, 2017. Aceito em 05 de Fevereiro, 2017.

Este trabalho apresenta resultados de uma revisão da literatura acerca da implementação da metodologia interativa de ensino Peer Instruction (PI). Respondemos às seguintes questões de pesquisa: Em quais contextos de ensino (nível de ensino, país, área de ensino e disciplinas) pesquisadores têm investigado o PI? Que impactos o PI tem produzido na aprendizagem dos estudantes? Quais são os resultados instrucionais da implementação do PI em termos das atitudes dos professores em relação à metodologia e das modificações feitas na estrutura original da metodologia? Quais são os principais aspectos teóricos e metodológicos que pesquisadores utilizam para investigar implementações do PI? Os resultados da literatura apontam que grande parte das publicações foram conduzidas em universidades norte americanas, em disciplinas da área STEM, com destaque à disciplina de Física. A adoção do PI apresenta impactos positivos na aprendizagem conceitual dos estudantes, na habilidade de resolução de problemas e no desempenho acadêmico. Desenvolve sentimentos positivos relacionados à aprendizagem dos conteúdos e à metodologia. Professores realizam modificações ao implementar o PI integrando-o com outras metodologias, demonstrando sua flexibilidade. A maioria dos estudos é apoiada por análises empíricas e estatísticas, mas não por estruturas conceituais ou referenciais teóricos. Esta lacuna apresenta oportunidades para contribuições futuras.

Palavras-chave: revisão da literatura; Peer Instruction; metodologias interativas de ensino.

This paper presents a review of literature on the implementation of the interactive student-centered teaching method Peer Instruction (PI). We answer the following research questions: In which teaching contexts (education level, country, teaching area and disciplines) have researchers investigated PI? What student outcomes are detailed in PI implementation studies? What are the instructional outcomes of PI adoption among instructors, in terms of teacher attitudes towards the methodology and modifications made to the original structure of the methodology? What are the theoretical and methodological approaches researchers use to study PI implementation? The results of the literature demonstrate that the large majority of publications on PI implementation result from studies conducted at North American universities, in the STEM fields, particularly within the discipline of Physics. PI implementation shows increases in the conceptual learning of students, problem-solving ability and academic performance. Develops students' positive feelings related to content learning and the teaching methodology. Instructors make changes to the implementation of the method and integrate it with other teaching methods, demonstrating the methodology's flexible nature. Most studies on PI implementation are supported by empirical and statistical analysis but are not guided by formal conceptual or theoretical frameworks. This gap presents opportunities for future contributions.

Keywords: literature review; Peer Instruction; interactive teaching methods.

1. Introdução

Apesar de não se tratar de um conceito absolutamente novo, a aprendizagem ativa tem adquirido grande atenção nas últimas décadas. Muito frequentemente, a expressão “aprendizagem ativa” é reco-

nhecida como uma abordagem para a sala de aula, onde atividades são programadas para engajar os estudantes durante todo o processo de aprendizagem, em contraste ao ensino tradicional, caracterizado por aulas eminentemente expositivas [1]. Professores e pesquisadores da área de ensino têm voltado esforços no desenvolvimento de novas metodologias de ensino que visam o engajamento dos estudan-

*Endereço de correspondência:
maykonmuller@cavg.ifsul.edu.br.

tes durante o processo de aprendizagem como, por exemplo: *One-Minute Paper* [2], *Think-Pair-Share* [3], *Problem-Based Learning* [4], *Team-Based Learning* [5], *Just-in-time Teaching* [6], *SCALE-UP* [7]. Essas metodologias não esgotam, de maneira alguma, as possibilidades disponíveis para os professores. Em trabalho recente, Henderson et al. [8] investigaram a adoção de, pelo menos, 24 metodologias ativas de aprendizagem por professores em disciplinas de Física universitária.

Entre as metodologias ativas de aprendizagem, o *Peer Instruction* (PI) possui um reconhecimento preeminente [9]. O *Peer Instruction* é uma metodologia de ensino ativa centrada no estudante (*active student-centered pedagogy*) desenvolvida no início da década de 90 do século passado pelo professor de Física Eric Mazur da Universidade de Harvard. Desde sua criação, o PI vem ganhando destaque internacional por sua capacidade de engajar ativamente os estudantes durante o processo de aprendizagem, de intensificar a aprendizagem conceitual, além de desenvolver habilidades sociais e cognitivas [10, 11]. Tais resultados podem ser encontrados em pesquisas realizadas nos mais diversos contextos sociais, disciplinas e níveis de ensino [10–20].

Passadas aproximadamente duas décadas de sua criação, é oportuno traçar um panorama da produção acadêmica referente à aplicação do PI em ambientes formais de ensino. Recentemente, Vickrey et al. [21] publicaram uma ampla revisão da literatura, que abarca estudos sobre a eficiência do PI e sobre as atitudes de estudantes e professores em relação a ele, identificam aspectos críticos da implementação do método, assim como uma breve síntese dos principais contextos acadêmicos em que o PI tem sido adotado. Tal revisão inclui publicações em revistas especializadas e em anais de eventos. Um dos pontos fortes dessa revisão é a discussão acerca de aspectos críticos da implementação do PI, associados a cada uma das suas etapas. A presente revisão, além de ter um escopo mais amplo, incluindo artigos publicados em língua portuguesa, aprofunda a discussão acerca dos principais impactos no desempenho dos estudantes. Ademais, analisa os referenciais teóricos e metodológicos utilizados nas publicações revisadas. Por fim, nossa busca por publicações referentes ao PI também não encontrou trabalhos de revisão publicados em língua portuguesa. Definimos como questões de pesquisas a serem respondidas através da revisão: *i. Em quais contextos de ensino (nível de ensino, país, área de ensino e disciplinas) pesquisadores têm investigado o PI? – ii. Que impactos o Peer Instruction tem produzido na aprendizagem*

dos estudantes? – iii. Quais são os resultados instrucionais da implementação do PI em termos das atitudes dos professores em relação à metodologia e das modificações feitas na estrutura original da metodologia? – iv. Quais são os principais aspectos teóricos e metodológicos que pesquisadores utilizam para investigar implementações do PI?

Apresentamos, na sequência, o escopo da revisão realizada, incluindo aspectos históricos relacionados ao PI. Após, a metodologia empregada, discutindo as categorias criadas para a análise dos artigos e, ao final, os resultados, as discussões dos resultados e as conclusões.

2. Escopo da revisão

2.1. Aspectos históricos do Peer Instruction

Inicialmente desenvolvido para o curso introdutório de Física ministrado pelo professor Eric Mazur, o *Peer Instruction* (PI) reúne elementos das pedagogias centradas nos aprendizes (*learner-centered teaching*), na qual o estudante possui papel central no processo de aprendizagem. O PI foi desenvolvido no ano de 1991 [22], quando Mazur tomou contato com o teste *Force Concept Inventory* (FCI) [23] sobre concepções de Mecânica. As pesquisas de Hestenes e colaboradores apontaram que os estudantes que entram em cursos introdutórios de Física apresentam muitas concepções de “senso comum” acerca de fenômenos relacionados à Mecânica, e que o ensino tradicional é pouco eficiente para os estudantes transporem tais concepções e crenças sobre a Física.

Entre os anos de 1984 e 1990, Mazur ensinava seus estudantes, em cursos introdutórios de Física, de maneira convencional, com aulas expositivas acompanhadas de demonstrações. De maneira geral, seus estudantes tinham um desempenho bom em resolução de problemas considerados difíceis e avaliavam as aulas de maneira positiva. Ao aplicar o FCI em uma turma de graduandos em Física de Harvard, Mazur percebeu alguns alertas. O primeiro veio por meio de uma pergunta feita por uma estudante: “Professor Mazur, como devo responder a essas questões? De acordo com o que o senhor ensinou ou conforme o meu jeito de *pensar* a respeito dessas coisas?”. Apesar do alerta, os estudantes tiveram um desempenho um pouco superior no FCI do que obtiveram em um exame de metade do semestre da disciplina [22].

Frente aos resultados contraditórios obtidos com o FCI, Mazur decidiu colocar nos exames questões qualitativas simples e questões quantitativas difíceis. Ao comparar o desempenho dos estudantes em pa-

res de questões qualitativas e quantitativas (sobre o mesmo conceito), foi possível perceber que grande parte dos estudantes conseguia resolver as questões quantitativas difíceis, mas não tinha um desempenho satisfatório nas questões qualitativas simples. As dificuldades apresentadas pelos estudantes expõem certas fragilidades do ensino de Ciências. Inicialmente, é perceptível que estudantes que memorizam algoritmos de resolução podem ter um ótimo desempenho em disciplinas de Física, apesar de não compreenderem os conceitos físicos subjacentes. Além disso, frente ao bom desempenho dos estudantes em exames tradicionais, o professor, mesmo tendo muita experiência no ensino, pode ter a falsa sensação de que seu trabalho é eficiente. De maneira correlata, os estudantes, ao terem sucesso com seus algoritmos de resolução de problemas, podem ter o sentimento que dominam o conteúdo abordado na disciplina. Contudo, ao se depararem com situações em que tais algoritmos não são eficazes, podem se sentir frustrados.

Decidido a modificar tal cenário, Mazur começou a buscar alternativas para suas aulas, cujo foco passou para a discussão dos conceitos subjacentes, sem que a habilidade de resolução de problemas fosse negligenciada. O resultado foi a criação do PI, cujos objetivos são explorar a interação entre os estudantes durante as aulas e focar a atenção dos estudantes nos conceitos fundamentais. A estrutura proposta por Mazur para o PI pode ser sintetizada nas seguintes etapas [22, 24]:

1. Uma curta apresentação oral sobre os elementos centrais de um dado conceito ou teoria é feita por cerca de 20 minutos.
2. Uma pergunta de múltipla escolha, geralmente conceitual, denominada Teste Conceitual, é colocada aos alunos sobre o conceito (teoria) apresentado na exposição oral.
3. Os alunos têm entre um e dois minutos para pensarem silenciosamente sobre a questão apresentada.
4. Os estudantes registram suas respostas individualmente e as mostram ao professor usando algum sistema de respostas (por ex., *clickers* ou *flashcards*).
5. De acordo com a distribuição de respostas, o professor pode passar para o passo seis (quando a frequência de acertos está entre 35% e 70%), ou diretamente para o passo nove (quando a frequência de acertos é superior a 70%).
6. Os alunos discutem a questão com seus colegas por um a dois minutos.

7. Os alunos registram sua resposta revisada e as mostram ao professor usando o mesmo sistema de respostas do passo 4.
8. O professor tem um retorno sobre as respostas dos alunos a partir das discussões e pode apresentar os resultados para os alunos.
9. O professor então explica a resposta da questão aos alunos e pode ou apresentar uma nova questão sobre o mesmo conceito ou passar ao próximo tópico da aula, voltando ao primeiro passo.

A estrutura proposta por Mazur para o PI é similar à metodologia interativa *Think-Pair-Share* (TPS) [3]. Durante uma aula, o professor apresenta uma questão de múltipla escolha, ou um problema, e os estudantes tem um tempo para desenvolver suas respostas individuais (*Think*). Na sequência, os alunos interagem com os colegas comparando as respostas encontradas (*Pair*) e, então, compartilham as respostas com o professor (*Share*).

Usualmente a votação (etapas 4 e 7) é feita por meio de algum sistema de resposta como *flashcards* (cartões de resposta) ou *clickers*, espécie de controles remotos individuais que se comunicam por radiofrequência com o computador do professor. Mais recentemente, sistemas de resposta envolvendo quaisquer dispositivos com acesso à internet, tais como *notebooks*, *smartphones* e *tablets* vêm se mostrando uma alternativa promissora, tanto por se valerem de aparelhos que os próprios estudantes já possuem, quanto por viabilizar o envio de respostas para questões abertas.

Desde a formulação inicial do PI, eram previstas tarefas prévias às aulas, nas quais uma ou mais seções de um livro-texto eram lidas pelos estudantes. Para garantir que a tarefa havia sido realizada, o professor aplicava *quizzes* no período de 20 min antes da aula até 5 min após seu início, valendo um percentual da nota final dos estudantes. Inicialmente, não havia o intuito de obter um *feedback* sobre as dificuldades do aluno, tampouco eram abertas possibilidades de perguntas sobre o conteúdo aos estudantes. O principal objetivo era garantir que os alunos, em algum momento, passassem pelos exemplos resolvidos dos livros-texto e acompanhassem os desenvolvimentos algébricos de equações que não mais seriam abordadas em aulas, pois foram retirados para abrir espaço para a discussão entre os alunos.

Com o passar do tempo, as tarefas de leituras foram potencializadas pelo uso do *Just-in-Time Teaching* (JiT) [6, 25], que trouxe como possibilidade adicional o professor conhecer as dificuldades mani-

festadas pelos alunos para preparar as suas aulas. Apesar do PI contar com tarefas de leituras prévias às aulas, convencionou-se, após seu uso conjunto com o JiTT, que o PI está relacionado com as atividades presenciais em sala de aula (i.e. breves exposições orais, aplicação de questões conceituais, votações, interações entre os pequenos grupos e explicações por parte do professor) e as atividades pré-classe estão relacionadas ao JiTT.

2.2. Metodologia

O levantamento dos artigos abarcados por esta revisão da literatura inclui pesquisas publicadas em revistas entre os anos de 1991 e 2015, e foi realizado recorrendo a duas bases de dados internacionais: *Social Sciences Citation Index* (SSCI) (através da plataforma *Web of Science*) e *Education Resources Information Center* (ERIC). A busca dos artigos foi estendida a periódicos específicos¹, classificados nos extratos A1 (19 periódicos), A2 (23 periódicos) e B1 (80 periódicos) na área de Ensino, da plataforma Sucupira (Qualis 2014). Na sequência, descrevemos os critérios utilizados, culminando na exclusão de alguns artigos e inclusão de outros.

No banco de dados ERIC, usamos como filtro no campo de busca o termo “Peer Instruction” contido em artigos em revistas com revisão pelos pares. Resultaram 198 artigos, dos quais 24 foram publicados antes de 1991 e por isso foram excluídos. Os 174 itens encontrados tiveram seus títulos e *abstracts* revisados manualmente em busca da palavra chave “Peer Instruction”, bem como de indícios de que o artigo tratava de uma pesquisa sobre o método. Dos 174 artigos, 94 não continham o termo Peer Instruction no título ou no *abstract*, um não era artigo de pesquisa e não tinha relação com PI e 16 artigos apresentavam o termo “Peer Instruction” como referência à interação entre os estudantes (*peer teaching, peer tutors, peer-to-peer instruction*) em estratégias didáticas interativas, mas não à metodologia PI especificadamente. Os 64 artigos restantes foram lidos integralmente.

Na base de dados SSCI, através da plataforma *Web of Science*, usamos como filtro o termo “Peer Instruction” no campo de tópico ou no título, contido em artigos publicados entre os anos de 1991 e 2015. Usamos, então, os filtros “articles” (n=131) em “document type” e “english” (n=130) em “language”. Os 130 itens encontrados tiveram seus títulos e *abstracts* revisados manualmente em busca da palavra

chave “Peer Instruction”, bem como de indícios de que o artigo tratava de uma pesquisa sobre o método. Após a leitura dos *abstracts*, 50 foram excluídos por não conterem o termo Peer Instruction, e 30 não foram identificados como pesquisas relacionadas ao método, pois utilizavam o termo “Peer Instruction” como referência à interação entre os estudantes (*peer teaching, peer tutors, peer-to-peer instruction*) e não à metodologia PI especificadamente; totalizando, dessa forma, 50 artigos que foram lidos integralmente. O total de artigos, obtidos através das plataformas SSCI e ERIC, lidos integralmente foi de 87 (haja vista que 27 artigos foram encontrados em ambas as plataformas de busca), sendo que destes, 19 foram excluídos, pois tratavam de pesquisa que: apenas citavam o PI [26–29]; abordavam estratégias didáticas que se baseiam na filosofia do “*peer teaching*” [30–36]; relata uso de sistemas de votação ou sistemas para apresentação de questões conceituais [37–39]. Além disso, outros dois artigos não foram encontrados através de meios eletrônicos para a leitura.

Nos periódicos classificados nos extratos A1, A2 e B1 na área de Ensino, utilizamos como filtro os termos “Peer Instruction”, “Instrução pelos Colegas”, “Instrução pelos Pares” e “Instrucción pelos colegas” e “Instrucción por Pares”. Através das buscas, encontramos um total de 13 artigos que se enquadraram no escopo de nossa revisão; destes 12, 8 foram encontrados através da busca nas bases de dados e um é um artigo de divulgação do PI [24]. Por conseguinte, nossa revisão abarca o número total de 72 artigos de pesquisa sobre o PI.

3. Resultados e Discussões

A apresentação dos resultados encontrados dar-se-á de acordo com as questões de pesquisa apresentadas anteriormente. Reservamos uma subseção para cada uma das perguntas.

3.1. Contexto de aplicação do PI

Os artigos compreendidos nesta revisão referem-se a estudos em diferentes níveis de ensino, a saber: Ensino Médio (n=5), Universidade do tipo “*College*” (n=7), Universidade (n=57) e Ensino Médio e Universidade (n=1). Os artigos de Vickrey et al. [21] e Araujo e Mazur [24] não foram contabilizados por se tratarem de um trabalho de revisão da literatura e de divulgação da metodologia, respectivamente. Há enorme predomínio no número de publicações cujo contexto de pesquisa é a universidade (aproximada-

¹A revisão foi realizada apenas nos periódicos disponíveis *on-line*.

mente 90%). A grande predominância de publicações no Ensino Superior pode ser interpretada de algumas maneiras. Em primeiro lugar, muitas pesquisas são realizadas onde os grupos de pesquisa em ensino estão situados. Dessa forma, o ambiente foco de pesquisa acaba sendo a instituição de nível superior. Além disso, as universidades enfrentam muitos problemas relacionados ao baixo nível de compreensão dos estudantes, ao índice de reprovação nas disciplinas introdutórias e a baixa motivação dos estudantes em aprender os conteúdos do curso. Adotar, portanto, novas metodologias de ensino, bem como compreender suas respectivas eficiências, é de extrema importância. Incontestavelmente, a Educação Básica também é afetada por problemas similares aos do Ensino Superior, especialmente no que se refere a motivação dos estudantes em aprenderem os conteúdos abordados em sala de aula. Frente o reduzido número de publicações, pesquisar os impactos do PI em ambientes formais de ensino básico é, portanto, uma linha de pesquisa promissora e necessária.

Em relação aos países em que essas pesquisas foram realizadas, conforme pode ser visto no Gráfico 1 há uma notória concentração de trabalhos na América do Norte, especificamente nos Estados Unidos da América (n=36), Canadá (n=5) e Estados Unidos e Canadá (n=1). O total de 42 artigos na América do Norte é superior à soma de todos os outros continentes juntos. Na África encontramos apenas um artigo [40]. Apesar dessa grande diferença no número de publicações, verifica-se que há um movimento de pesquisadores na busca de resultados da implementação do PI em contextos diferentes daquele em que foi criado, ou seja, em universidades americanas.

As áreas do conhecimento na qual os artigos foram desenvolvidos encontram-se no Gráfico 2. Nota-se uma predominância de artigos publicados na área

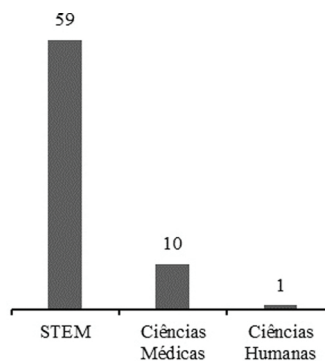


Gráfico 2: Distribuição de publicações por área do conhecimento.

de STEM², com aproximadamente 85% dos artigos³. Dentro da grande área STEM, a Física, com 31 artigos, é a disciplina em que o PI foi mais investigado, seguido pela Matemática (n=7) e pela Engenharia (n=7). Novamente, percebemos que o PI foi mais investigado na disciplina em que foi criado, a Física, apesar de que acabou se estendendo à toda grande área STEM.

3.2. Impactos causados pelo PI na aprendizagem

Nessa categoria, só incluímos os artigos que possuem questões de pesquisa e/ou objetivos de pesquisa explícitos, contendo a descrição da metodologia de pesquisa utilizada e, em alguns casos, o referencial teórico adotado. Abordaremos estudos sobre os impactos, proporcionados pela adoção do PI, na aprendizagem conceitual, na resolução de problemas, em diferentes gêneros e etnias, desempenho nas disciplinas, crenças, atitudes e motivação dos estudantes, bem como na retenção dos estudantes (nas disciplinas e nos cursos). Pesquisas cujo objetivo era compreender a importância dos elementos que constituem a estrutura do PI também são apresentadas nessa seção.

3.2.1. Aprendizagem Conceitual

Grande porção dos artigos (n=36) publicados sobre a adoção do PI possuía objetivos e/ou questões de pesquisa intimamente ligadas à análise do impacto

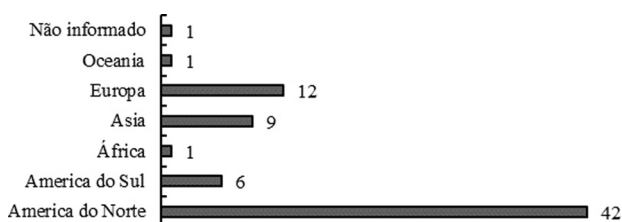


Gráfico 1: Distribuição de publicações por continente.

² A sigla STEM é um acrônimo para as disciplinas de Ciências, Tecnologia, Engenharia e Matemática (*Science, Technology, Engineering e Mathematics*). Os cursos da grande área STEM incluem, entre outros, Física, Astronomia, Química, Biologia, Engenharias, Ciência da Computação e Matemática.

³ Aqui também não foram categorizados os artigos de Vickrey et al. [21] e Araujo e Mazur [24].

na aprendizagem conceitual dos estudantes, apesar de não haver, em muitos casos, clareza quanto à aceção atribuída ao termo “aprendizagem conceitual”. Categorizamos, nesta seção, todos os artigos que avaliaram a aprendizagem dos estudantes por meio de questionários padronizados, questões conceituais utilizadas durante as aulas, ou até mesmo testes desenvolvidos pelos próprios autores.

Optamos por agrupar os artigos por área de conhecimento, *STEM*, *Ciências Médicas*, *Ciências Humanas*, e, dentro de cada uma dessas áreas, agrupamos pelos instrumentos de coleta de dados utilizados.

3.2.1.1. STEM

Aproximadamente 50% dos artigos (n=31) da grande área STEM (n=59) tiveram como objetivo investigar os impactos, proporcionados pela adoção do PI, na aprendizagem conceitual dos estudantes, cuja melhoria é vista como um dos principais resultados de sua implementação. Frequentemente, o PI é comparado às aulas tradicionais e, a subsequente avaliação da aprendizagem se dá através de testes padronizados (e.g. *Force Concept Inventory* – FCI [23], *Conceptual Survey in Electricity and Magnetism* – CSEM [41], *Geoscience Concept Inventory* – GCI [42]). Em algumas ocasiões, outros instrumentos, tais como testes construídos pelos pesquisadores e questões conceituais, utilizadas durante as aulas, são empregados para subsidiar as investigações.

A metodologia de pesquisa empregada inclui, em todos os artigos, análises quantitativas. Usualmente, o delineamento experimental envolve um grupo de controle (aulas tradicionais) e um grupo experimental (PI ou modificação). Em alguns estudos, um teste padronizado é aplicado antes e depois da implementação do PI [10, 11, 13, 14, 20, 43]. Nesses casos, o tratamento dos dados se dá através de testes de comparação (e.g. *independent t-test*, *ANOVA*, *Wilcoxon test*) para avaliar se há diferenças estatisticamente significativas entre o grupo experimental e o grupo de controle, correlações entre o desempenho no teste padronizado e outras variáveis independentes (e.g. quantidade de horas dedicadas pelo estudante para estudar) [11], etc. Outra técnica muito utilizada é o cálculo do ganho normalizado definido por Hake [44], como

$$g = \frac{\%pós - \%pré}{100\% - \%pré}, \quad (1)$$

onde $\%pré$ corresponde ao escore obtido no teste antes da implementação e $\%pós$, após a implementação.

O ganho normalizado mede a melhora no escore do estudante em um teste, levando em consideração seu próprio desempenho na primeira e na segunda aplicação do teste. O numerador da Equação 1 corresponde ao ganho efetivo obtido pelo estudante; o denominador, à melhora máxima possível de ser alcançada por aquele estudante. O valor de g pode variar entre 0 e 1 (ou entre 0% e 100%), sendo que resultados mais próximos de 1 correspondem a uma melhora mais acentuada. Os resultados negativos, obtidos quando o estudante apresenta um escore superior na primeira aplicação do teste, são desconsiderados da análise. O ganho normalizado também pode ser calculado com os escores dos Testes Conceituais; nesse caso, o numerador da Equação 1 é determinado pela diferença entre a percentagem de acertos após e antes da discussão entre os colegas. O denominador corresponde, portanto, a máxima melhora na percentagem de acertos para a questão conceitual.

3.2.1.1.1. Testes envolvendo questões conceituais

Devido à ausência de testes padronizados para determinados conteúdos e disciplinas, uma parcela dos pesquisadores avalia a aprendizagem dos estudantes por meio da aplicação de testes construídos especialmente para sua pesquisa [40, 45–51].

Os estudos de Van Dijk et al. [46] e Schmidt [48], conduzidos em disciplinas para estudantes de Engenharia, mostram que o PI proporcionou uma melhora na aprendizagem conceitual de conceitos de Mecânica e Dinâmica, respectivamente. Em específico, Van Dijk et al. [46] compararam o desempenho dos estudantes em três situações distintas: aulas tradicionais; com o PI; e uso de sistemas de votação (sem PI). Na análise do pré-teste não foram encontradas diferenças estatisticamente significativas entre os grupos. Os resultados do pós-teste indicam que os estudantes que tiveram aulas com o PI tiveram um escore superior no teste, sendo a diferença nos escores estatisticamente significativa com $p = 0,025$ (tradicional: $4,82 \pm 1,5$; sistema de votação: $4,00 \pm 1,71$; PI: $5,23 \pm 1,31$). Já Schmidt [48] comparou a aprendizagem conceitual dos estudantes que tiveram aulas com o PI (grupo experimental) com uma turma de controle (ensino tradicional). A análise estatística do pré-teste mostrou que as duas turmas em que o PI foi adotado e a turma de controle (que teve aulas tradicionais) não possuíam diferenças estatísticas de conhecimento prévio. O ganho normalizado das duas turmas experimentais foi superior ao da turma de controle ($\langle g \rangle_{PI-1} = 0,46$; $\langle g \rangle_{PI-2} = 0,59$; $\langle g \rangle_{Controle} = 0,36$).

A melhora na aprendizagem conceitual dos estudantes, medida com testes desenvolvidos pelos pesquisadores, também foi encontrada em cursos de Biologia [49], Matemática [40, 45], Física [47, 50] e Química [51]. Um resultado interessante foi obtido por Jones et al. [49], que mostraram que o PI ampliou a capacidade de estudantes de um curso de Entomologia em aplicar o conhecimento conceitual a situações do mundo real. Em termos do desempenho dos estudantes nos testes aplicados, Pizer [45], Cox e Junkin [47], Zhu e Singh [50], Morice et al. [51] e Ouko et al. [40] dão evidências que os estudantes que tiveram aulas com o PI tendem a obter um escore superior na segunda aplicação do teste (pós-teste), em comparação com os escores de estudantes com aulas tradicionais.

3.2.1.1.2. Testes Conceituais

Testes Conceituais foram utilizados como fonte de evidências dos impactos proporcionado pela adoção do PI em uma parcela dos trabalhos encontrados nesta revisão ($n=9$). Para mensurar a influência da discussão entre os colegas na aprendizagem dos estudantes, as pesquisas avaliam diferentes aspectos: a qualidade das justificativas para os Testes Conceituais, a convergência para a resposta correta, o ganho normalizado para os Testes Conceituais, e o desempenho em questões isomórficas. Passamos a discutir os resultados referentes a cada um desses aspectos.

Brook e Koretsky [52] e Koretsky e Brooks [53], em estudos realizados em turmas de Química e Engenharia Química, respectivamente, mostram que a qualidade das explicações para as respostas melhora após a interação entre colegas. Outro elemento de destaque, encontrado por Brook e Koretsky [52], é um aumento no nível de confiança dos estudantes em relação às suas respostas, após a etapa de interação com os colegas. O nível de confiança dos estudantes foi medido através de uma escala tipo *Likert* que variava entre 1 (substancialmente inseguro) a 5 (substancialmente seguro). Em cada rodada de votação, os estudantes deviam optar pela resposta que julgavam correta e pelo grau de confiança que possuíam em relação a sua resposta.

A convergência para resposta correta, após a etapa de discussão entre os colegas, é outra evidência amplamente divulgada como um dos pilares do PI. Como é bem sabido, no PI após a primeira rodada de votações, caso haja uma dispersão nas respostas, o professor sugere que os estudantes encontrem colegas que optaram por respostas diferentes das suas para discutir acerca da questão. Os resultados indicam que frequentemente os estudantes, ao inte-

ragirem uns com os outros, acabam por encontrar a resposta correta para o Teste Conceitual sem que haja, necessariamente, uma intervenção do professor [10, 52, 54–57].

Outra análise conduzida com a distribuição de acertos e erros dos estudantes nas duas etapas de votação é o ganho normalizado. Lee et al. [56], ao investigarem a eficácia da adoção do PI em seis disciplinas de cursos da Ciência da Computação, duas de Arquitetura de Computadores e quatro de Teoria da Computação, mostraram que a aprendizagem conceitual foi intensificada devido à etapa de discussão entre os colegas, com ganhos normalizados entre 34% e 45% (média de 39%). Michinov et al. [58] avaliaram uma modificação do PI que incluía a técnica denominada “*stepladder*”, com a versão clássica do PI e instrução individual (os estudantes respondiam aos Testes Conceituais de maneira individual), em uma disciplina de Química para alunos do curso de Engenharia Química. Os alunos que tiveram aulas com a versão modificada do PI obtiveram um ganho normalizado para os Testes Conceituais, classificados como difíceis, superior aos outros dois grupos (versão modificada $\langle g \rangle = 0,80 \pm 0,29$; versão clássica $\langle g \rangle = 0,53 \pm 0,31$; individual $\langle g \rangle = 0,23 \pm 0,33$).

A aprendizagem dos estudantes também foi aferida com a aplicação de questões denominadas “isomórficas⁴”. Nesta revisão, verificamos que Zingaro e Porter [59], inspirados em Porter et al. [60] e Smith et al. [61], buscaram avaliar a aprendizagem dos estudantes em uma disciplina de introdução à computação através da análise das respostas a questões isomórficas com e sem a intervenção do professor. A intervenção do professor ocorreu antes de os estudantes responderem às questões isomórficas, diferenciando-se, dessa forma, dos trabalhos de Porter et al. [60] e de Smith et al. [61]. Os autores argumentam que o papel do professor é extremamente fundamental para a boa compreensão dos conceitos envolvidos nas questões conceituais. Os estudantes que após a segunda rodada de votação receberam uma explicação do professor acerca da questão, tiveram um desempenho superior aos estudantes que não tiveram a intervenção do professor (84% de acertos *vs.* 76% de acertos, respectivamente). Tal resultado foi mais expressivo em questões que os autores avaliaram como difíceis.

3.2.1.1.3. Testes Padronizados

Testes Padronizados se constituíram na fonte de dados mais frequente em publicações da área STEM,

⁴ Questões Isomórficas são questões conceituais que apresentam situações diferentes mas abordam o mesmo conceito [61].

representando um total de 14 publicações. De maneira geral, as pesquisas que utilizaram testes padronizados avaliaram os impactos na aprendizagem dos estudantes que tiveram aulas com o PI (turma experimental) em comparação com estudantes que tiveram aulas tradicionais (turma de controle).

Crouch e Mazur [10], de Barros et al. [20], Lasry et al. [14] e Rudolph et al. [11] realizaram pesquisas em disciplinas introdutórias de Física, que abordam tópicos de Mecânica. Através da aplicação do teste padronizado *Force Concept Inventory* (FCI) [23], no formato de pré-teste/pós-teste, as investigações de Crouch e Mazur [10] (realizadas em uma universidade de ponta) e Lasry et al. [14] (realizado em uma universidade do tipo *two-year college*) encontraram um desempenho superior, em termos do ganho normalizado, dos estudantes que tiveram aulas com PI, em comparação aos grupos de controle. Esse resultado também foi encontrado por de Barros et al. [20], em estudo realizado em uma universidade brasileira. Já Rudolph et al. [11] não encontraram diferenças estatísticas entre o grupo experimental e o grupo de controle na versão completa do FCI. No entanto, ao avaliar apenas as questões que versavam sobre os conceitos trabalhados durante o estudo, os autores encontraram um ganho normalizado superior para o grupo experimental.

A aprendizagem de conceitos de Eletromagnetismo e Magnetismo foi avaliada por Meltzer e Manivannan [62], Lenaerts et al. [13], Gok [43] e Rudolph et al. [11]. Como fonte de dados, foram utilizados os testes padronizados *Magnetism Concept Inventory* (MCI) [41], *Conceptual Survey in Electricity* (CSE) [41] e *Conceptual Survey in Electricity and Magnetism* (CSEM) [41]. Gok [43] e Rudolph et al. [11] mostram que os estudantes que tiveram aulas com o PI obtiveram ganhos normalizados superiores, em comparação aos estudantes dos grupos de controle, no teste CSEM. Também Lenaerts et al. [13], através da análise do ganho normalizado dos escores do MCI, concluíram que os estudantes que tiveram aulas com o PI tiveram uma aprendizagem superior dos conceitos tratados na disciplina (grupo experimental: $\langle g \rangle = 0,59 \pm 0,10$; grupo de controle: $\langle g \rangle = 0,39 \pm 0,08$). Meltzer e Manivannan [62] utilizaram dois testes padronizados, o CSE e o CSEM (apenas questões sobre eletricidade), e um teste sobre circuitos elétricos (retirados de Shaffer e McDermott [63, 64]). Os resultados com ambos os testes padronizados mostram que os estudantes que tiveram aulas com o PI tiveram ganhos normalizados superiores aos que tiveram aulas tradicionais (CSE – grupos de controle: $\langle g \rangle = 0,46$ e $\langle g \rangle = 0,53$; CSE – grupos ex-

perimentais: ganhos entre $\langle g \rangle = 0,64$ e $\langle g \rangle = 0,69$; CSEM – grupos experimentais: $\langle g \rangle = 0,22$; CSEM – grupos experimentais: ganho entre $\langle g \rangle = 0,64$ e $\langle g \rangle = 0,71$). Os estudantes das turmas em que o PI foi adotado também tiveram um desempenho satisfatório no teste sobre circuitos elétricos, quando comparados aos resultados de Shaffer e McDermott [64].

Cox e Junkin [47] buscaram investigar o impacto da adoção de uma versão modificada do PI na aprendizagem de conceitos de Ótica e Cinemática em turmas de laboratório de Física. A principal mudança feita na estrutura do método foi a etapa de votação, em que os alunos votavam em grupo, e não individualmente como originalmente proposto por Mazur [22]. Através da aplicação de uma versão reduzida do “*Force and Motion Conceptual Evaluation*” (FMCE) (apenas oito questões), os autores concluíram que o PI é uma metodologia eficiente para a aprendizagem de conceitos de Física em aulas de laboratório. Os resultados mostram que o ganho normalizado médio da turma de controle ($\langle g \rangle = 0,25$), no FMCE, foi inferior ao da turma experimental ($\langle g \rangle = 0,5$).

Algumas investigações avaliaram o impacto, na aprendizagem dos estudantes, do uso conjunto do PI com outras metodologias interativas. Busch [65] avaliou a aprendizagem de conceitos de Mecânica, em uma disciplina que aborda diversos conceitos relacionados às Ciências Ambientais, de estudantes que tiveram aulas com o uso conjunto do PI o SCALE-UP [7] e o *Just-in-time Teaching* [6]. O ganho normalizado no FCI obtido pela turma experimental foi superior ao da turma de controle ($\langle g \rangle = 0,4$ vs. $\langle g \rangle = 0,25$)

O uso em conjunto do PI com a metodologia denominada *Lectures Tutorials* (LT) foi investigado por Gray e Steer [66]. Ao analisarem os escores obtidos pelos estudantes no teste padronizado *Geoscience Concept Inventory* (GCI) [42], os autores concluíram que o uso do PI combinado com a LT não promove ganhos estatisticamente significativos na aprendizagem de conceitos de geociências quando comparados ao uso exclusivo da LT. Contudo, os estudantes que tiveram aulas com o uso combinado das metodologias tiveram um escore no pós-teste superior ao grupo que utilizou apenas os LT ($42,7 \pm 2,8\%$ vs. $39,9 \pm 2,4\%$)

Cahyadi [67] avaliou o uso do PI combinado com outras metodologias interativas (*activer learning problems sets, constructivist classroom dialogue; demonstrations*) em uma disciplina introdutória de Física para dois cursos de engenharia em uma universidade na Indonésia. A análise do ganho norma-

lizado mostrou que as duas turmas em que o PI foi adotado em conjunto com outras metodologias interativas tiveram um ganho normalizado superior a três das quatro turmas em que aulas tradicionais foram ministradas ($\langle g \rangle = 0,45$ e $\langle g \rangle = 0,31$ vs. $\langle g \rangle = 0,17$, $\langle g \rangle = 0,12$, $\langle g \rangle = 0,31$). Em uma das turmas com aulas tradicionais o ganho normalizado foi superior ($\langle g \rangle = 0,50$), mas o instrutor admitiu ter “ensinado os estudantes para o teste”.

Os impactos na aprendizagem conceitual dos estudantes proporcionados pelo PI também foram contrastados com outras metodologias interativas de ensino. Kalman et al [68] buscaram avaliar a eficiência do PI em comparação com a metodologia denominada Grupos Colaborativos em uma disciplina introdutória de Física. Não houve diferenças estatisticamente significativas entre as duas metodologias de acordo com a análise dos escores obtidos pelos estudantes no FCI, usado como pré e pós teste.

Nessa mesma linha de investigação, Keiner e Burns [69] analisaram a adoção de duas metodologias interativas, o PI e o SCALE-UP [7], em turmas de Física de uma universidade americana. A adoção de ambas as metodologias foi comparada com uma turma de controle, com aulas tradicionais. O FCI foi utilizado como pré-teste e pós-teste. Os resultados obtidos pela turma de controle foram comparados com os resultados de quatro turmas em que o PI foi empregado, bem como com sete turmas em que o SCALE-UP foi adotado. A análise dos escores no FCI obtidos pelos estudantes mostram que ambas as metodologias interativas são mais eficientes que a abordagem tradicional, sendo o ganho normalizado médio da turma de controle de $\langle g \rangle = 0,18$ e das turmas experimentais $\langle g \rangle = 0,24$ (PI) e $\langle g \rangle = 0,36$ (SCALE-UP).

3.2.1.2. Ciências Médicas

A aprendizagem conceitual também foi tema de pesquisa na área de Ciências Médicas. Rao e DiCarlo [70] avaliaram o impacto da interação entre os colegas no desempenho de estudantes em questões conceituais em um curso de Fisiologia para estudantes de Medicina. As questões respondidas pelos estudantes foram categorizadas em: revisão; compreensão, aplicação e habilidades analíticas; e habilidades de síntese e avaliação. Nas três categorias ocorreu um aumento estatisticamente significativo ($p < 0,05$) no escore dos estudantes após a discussão entre os colegas (revisão: de $94 \pm 1,8\%$ para $99,4 \pm 0,4\%$; compreensão, aplicação e habilidades analíticas: de

$82,5 \pm 6\%$ para $99,1 \pm 0,9\%$; habilidades de síntese e avaliação: $73,6 \pm 11,6\%$ para $99,8 \pm 0,24\%$).

Nessa mesma linha de investigação, a pesquisa conduzida por Cortright et al. [71] avaliou a aprendizagem conceitual dos estudantes de um curso de Fisiologia do Exercício em uma universidade americana, por meio de Testes Conceituais utilizados em aula. O grupo dos estudantes que respondeu às questões conceituais e interagiram com os colegas, conforme a estrutura do PI sugere, teve um desempenho superior. Sem o PI, os estudantes responderam corretamente em $44 \pm 5\%$ do tempo. Já os estudantes que tiveram o PI, responderam corretamente $59 \pm 6\%$ do tempo, sendo a diferença entre os grupos estatisticamente significativa com $p = 0,02$.

Por outro lado, Petersen et al. [72] focaram a aprendizagem do diagnóstico de distúrbios ácido-base através da análise de gás do sangue arterial. Os autores encontraram resultados positivos tanto para o grupo experimental (PI) como para o grupo de controle (atividades colaborativas em grupo), revelando que ambas as metodologias podem contribuir para a aprendizagem dos estudantes. Passados dois anos desse estudo, Hartmann et al. [73] avaliaram a retenção e o grau de confiança dos estudantes acerca dos mesmos conteúdos investigados por Petersen et al. [72]. Os autores [73] concluíram que os estudantes que tiveram aulas com atividades colaborativas e com o PI obtiveram uma percentagem de retenção maior que estudantes de um grupo de controle, que tiveram aulas tradicionais.

Schuller et al. [74] avaliaram a aprendizagem conceitual de estudantes residentes em cirurgia geral de uma universidade americana através de Testes Conceituais aplicados em sessões de revisão do conteúdo. Os autores compararam os resultados entre os estudantes que tiveram aulas com a combinação entre PI e JiTT, com os estudantes que tiveram aulas tradicionais. Os resultados encontrados pelos autores mostram que os estudantes que tiveram aula com o PI e o JiTT obtiveram um índice de acertos superior ($p < 0,01$) ao restante dos estudantes.

3.2.1.3. Ciências Humanas

Entre os artigos abarcados por esta revisão, apenas a pesquisa de Butchart et al. [15] investigou os impactos na aprendizagem conceitual de estudantes na grande área de conhecimento Ciências Humanas, na aprendizagem de estudantes de Filosofia através do teste padronizado *California Critical Thinking Skills Test* (CCTST) [75]. Os estudantes que tiveram aulas com o PI obtiveram um ganho

estatisticamente significativo nos escores do CCTST ($17,23\% \pm 8,5\%$), em comparação aos estudantes que tiveram aulas tradicionais ($7,85\% \pm 5,5\%$). Esses resultados foram também suportados pelas atitudes positivas dos estudantes em relação à contribuição do PI para a compreensão dos conceitos abordados na disciplina.

3.2.2. Resolução de Problemas

Apesar do grande destaque dado aos impactos do PI na aprendizagem conceitual dos estudantes, também é amplamente discutida na literatura a contribuição do PI na habilidade de resolução de problemas. Usualmente, o ensino tradicional acaba por dar grande atenção à resolução de problemas quantitativos durante as aulas, sendo os problemas “mais relevantes” tratados de maneira minuciosa pelo professor. Pela tradição, pode-se ter a impressão que essa é a maneira mais adequada de desenvolver habilidades de resolução de problemas nos estudantes. Sendo assim, à primeira vista, a estrutura do PI, por focar a atenção dos estudantes na compreensão de conceitos, não favorece a construção de habilidades dos estudantes na resolução problemas quantitativos.

Localizamos no total oito estudos sobre a contribuição do PI no desenvolvimento de habilidades de resolução de problemas quantitativos e qualitativos. Alguns deles focam na resolução de problemas de Mecânica [10, 14, 67, 76], de Fisiologia [71, 77] e de Eletricidade [43, 62]. Como fonte de dados, as pesquisas valeram-se de problemas quantitativos contidos nos exames de final da disciplina, em testes desenvolvidos pelos autores, em um teste padronizado denominado *Mechanics Baseline Test* (MBT) [78], bem como em problemas resolvidos durante as aulas. Em todos os estudos foram conduzidas análises quantitativas e, em alguns casos, buscou-se diferenças estatísticas entre grupos experimentais e de controle.

Os impactos do PI na performance dos estudantes na resolução de problemas quantitativos de Mecânica foram mensurados por Gok [76], Lasry et al. [14], Cahyadi [67] e Crouch e Mazur [10], e de Eletricidade por Meltzer e Manivannan [62]. Lasry et al. [14], em um estudo conduzido em uma universidade do tipo “two-year college”, avaliaram o desempenho em resolução de problemas dos estudantes em teste aplicado ao final do curso. O teste era composto por 90% de problemas quantitativos de Mecânica. Os estudantes que tiveram aulas com o PI tiveram um escore de 68% de acertos, já os estudantes que tiveram aulas tradicionais obtiveram um escore de 63%

de acertos. Apesar da disparidade entre os resultados, segundo os autores, não foram encontradas diferenças estatisticamente significativas entre os escores. Crouch e Mazur [10] através da aplicação do MBT [78] ao final de disciplinas introdutórias de Física, investigaram a adoção do PI durante 10 anos em Harvard. Os estudantes das turmas em que o PI foi utilizado tiveram um escore superior nas questões quantitativas do MBT, variando entre 66% e 73%, que a turma experimental, cujo escore foi de 62%.

Gok [76], através da análise estatística de um teste desenvolvido pelo autor, composto de 20 problemas quantitativos, de múltipla escolha, relacionados com a aplicação das Leis de Newton, mostrou que o grupo experimental (aulas com PI), obteve um desempenho superior ao grupo de controle (aulas tradicionais), quando comparadas as médias do pós-teste (experimental: 69,07%; controle: 60,57%; $p < 0,05$) e os ganhos normalizados (experimental: $\langle g \rangle = 0,61$; controle: $\langle g \rangle = 0,51$). Meltzer e Manivannan [62] também reportaram resultados positivos quanto adoção do PI favorece ao desenvolvimento de habilidade de resolução de problemas de Eletricidade. Os autores valeram-se da análise do desempenho dos estudantes em questões específicas adicionadas ao exame de final de curso. Os estudantes que tiveram aulas com o PI tiveram, em média, um escore 20% superior aos estudantes que tiveram aulas tradicionais. Já Cahyadi [67], ao analisar os escores de um exame envolvendo problemas quantitativos de Mecânica, não encontrou diferenças significativas entre os estudantes que fizeram uso combinado do PI com outras metodologias interativas de ensino e os estudantes que tiveram aulas tradicionais.

Em relação ao desenvolvimento de habilidades específicas relacionadas ao processo de resolução de problemas, destacamos a pesquisa conduzida por Gok [43]. Os resultados, obtidos através da aplicação de um teste ao final da disciplina, contendo cinco problemas, apontam que os estudantes que tiveram aulas com o PI tiveram um desempenho superior em comparação aos estudantes que tiveram aulas tradicionais. A análise levou em consideração as três etapas de resolução de um problema: (1) identificação do princípio fundamental, (2) resolução e (3) revisão. Em todas as três etapas os estudantes que tiveram aulas com o PI tiveram um desempenho superior. Gok [76] desenvolveu uma sequência de questões para investigar a influência da interação entre os estudantes nas habilidades de resolução de problemas. Inspirado por Gok [43], as questões foram subdivididas em três categorias relacionadas

com as etapas de resolução de um problema: i. identificando os princípios fundamentais; ii. resolução; iii. verificação. O desempenho dos estudantes após a etapa de discussão entre os colegas foi superior nas três categorias. O grupo de controle também respondeu tais questões durante as aulas de maneira individual (sem interação entre os colegas), e o desempenho dos estudantes foi inferior ao grupo experimental após a etapa de discussão entre os colegas.

Cortright et al. [71] avaliou o impacto do PI na habilidade dos estudantes em resolverem problemas inéditos (“*novel problems*”) em um curso de Fisiologia do Exercício. Os estudantes que interagiram com os colegas durante a etapa de resolução de problemas obtiveram um escore superior aos estudantes que resolveram os problemas no formato tradicional ($47 \pm 5\%$ vs. $24 \pm 2\%$). A diferença, nesses resultados, foi estatisticamente significativa com $p < 0,05$. Giuliadori et al. (2006) [77], em um estudo desenvolvido em uma disciplina de Fisiologia, buscaram compreender se a habilidade dos estudantes em resolver problemas qualitativos é influenciada pela adoção do PI. Através da análise das respostas dos estudantes a problemas qualitativos de múltipla escolha, os autores puderam avaliar que os estudantes que engajaram-se em discussões com os colegas obtiveram um índice de acertos superior aos estudantes que apenas responderam individualmente as mesmas questões ($80.3 \pm 0.4\%$ vs. $59.3 \pm 0.5\%$).

3.2.3. Gênero e Etnia

As pesquisas categorizadas nessa seção buscam, de maneira resumida, averiguar possíveis diferenças nas maneiras como os diferentes gêneros e etnias são impactados pela adoção do PI. De maneira geral, as investigações não encontraram diferenças entre gêneros e etnias nos impactos proporcionados pelo PI [79–81]. Miller et al. [79], em uma pesquisa conduzida em 17 pequenas turmas de Cálculo, identificaram que o PI impacta positivamente no desempenho dos estudantes no exame final da disciplina, independentemente de gênero ou etnia, quando em comparação aos estudantes que tiveram aulas tradicionais.

Steer et al. [80] mostram que não existem diferenças de gênero nas distribuições das respostas dos estudantes nas duas etapas de votação previstas pela estrutura do PI. Além disso, a distribuição das respostas incorretas na primeira votação e corretas na segunda votação (categorizadas como incorreta-correta), quando analisada em termos das etnias e gênero, mostra que todos os grupos se beneficiam

da interação entre os colegas. Outros resultados encontrados apontam que os estudantes com o maior índice de acertos nos Testes Conceituais, em todos os grupos, tiveram um desempenho superior no pós-teste e na nota final de curso.

Investigação sobre o tempo gasto por homens e mulheres para registrar as suas respostas quando respondem aos Testes Conceituais, antes e após a discussão entre os colegas, foi realizada por Richardson e O’Shea [81]. A análise dos dados coletados em duas turmas de um curso introdutório de Física não evidenciou diferenças de gênero. De maneira geral, o PI permitiu que ambos os gêneros respondessem correta e rapidamente, enfatizando, dessa forma, o valor das metodologias ativas para a sala de aula.

3.2.4. Desempenho Acadêmico

O engajamento proporcionado pelos PI torna os estudantes mais responsáveis por sua aprendizagem ao longo das disciplinas. Uma das consequências imediatas desse aumento na responsabilidade dos estudantes é a melhora na performance em exames [82], bem como na disciplina como um todo [79, 83]. O desempenho acadêmico também apresenta melhoras quando o PI é utilizado com outras metodologias interativas de ensino, como é o caso dos “*Lecture Tutorials*” [66]. Além disso, a importância da interação entre os colegas na etapa de votação parece ser um fator determinante para o sucesso nos exames durante a disciplina [79]. Por outro lado, os resultados encontrados por Zingaro [84] apontam que o PI não favoreceu o desempenho acadêmico de estudantes em uma disciplina introdutória de um curso de Ciências da Computação. As notas dos exames dos estudantes que tiveram aulas com PI não tiveram diferenças estatisticamente significativas com as notas dos estudantes que tiveram aulas tradicionais.

A melhora na performance acadêmica, segundo Roth [85] e Arnesen et al. [86], beneficiou positivamente os estudantes que usualmente apresentam um baixo desempenho na disciplina, haja vista que tais estudos encontraram uma redução de conceitos considerados baixos após a adoção do PI. Arteaga e Vinken [83] mostraram que a assiduidade dos estudantes aumenta quando as aulas são ministradas com o PI.

A análise da correlação entre o desempenho dos estudantes nos Testes Conceituais e os escores dos estudantes nos exames, durante o semestre, foram realizados pelos trabalhos de Lucas [54], Lucas [87] e Morgan e Wakefield [55]. Lucas [54] encontrou uma

correlação positiva entre as notas dos estudantes e o índice de acertos nos testes conceituais. Com resultados similares, Lucas [87] encontrou uma correlação positiva entre o desempenho dos estudantes nas questões prévias às aulas, o desempenho nas questões conceituais em classe e a nota final do curso. Já Morgan e Wakefield [55] não encontraram correlação entre o desempenho dos estudantes na disciplina e na segunda rodada de votação dos testes conceituais respondidos durante as aulas. Os autores esperavam encontrar uma correlação entre a convergência para a resposta correta e a nota final da disciplina, o que não foi encontrado. Contudo, os autores encontraram uma correlação moderada entre o desempenho dos estudantes na primeira rodada de votações (antes da interação entre os estudantes) com a nota final do curso – os estudantes que acertaram nas duas rodadas de votação obtiveram conceitos superiores na disciplina.

3.2.5. Retenção

A taxa de estudantes que desistem das disciplinas ao longo do semestre é um fator preocupante para muitos professores, tanto em nível universitário quanto em nível de educação básica. Para evitar a evasão, a adoção de novas propostas de ensino deve contribuir para a criação de um ambiente receptivo e acolhedor, que promova nos estudantes o desenvolvimento de um sentimento positivo frente à disciplina e, ao engajarem-se nas atividades propostas, persistam até o final da disciplina. Nesse sentido, Lasry et al. [14] buscaram analisar o número de estudantes que finalizaram a disciplina em que aplicaram o PI, comparativamente a turmas com aulas tradicionais. Menos de 5% dos estudantes que tiveram aulas com o PI não finalizaram o curso. Em contraste, o índice de estudantes que desistiu do curso na turma experimental foi de 20,5%. Tais resultados apontam que a adoção do PI pode contribuir para a diminuição no número de alunos que desistem ao longo das disciplinas.

Conforme já discutido, o ensino tradicional apresenta resultados pouco positivos para a aprendizagem conceitual dos estudantes. Outro resultado é a baixa retenção dos estudantes nos cursos em que pretendem se graduar. De maneira geral, as disciplinas introdutórias de Ciências, por exemplo, tornam-se um divisor de águas nas vidas dos estudantes. As reclamações do chamado ensino tradicional incluem uma baixa qualidade na aprendizagem, levando muitos estudantes a mudarem de curso devido à baixa compreensão dos conceitos básicos [88]. O estudo de Watkins e Mazur [89] avaliou a retenção de es-

tudantes nos seus cursos *STEM* após terem tido contato com o PI em uma disciplina introdutória de Física. A porção de estudantes que tiveram aulas tradicionais e mudaram de curso foi duas vezes maior do que a porção de estudantes que tiveram aulas com o PI e trocaram de curso. Tais resultados são mantidos quando analisados por gênero.

3.2.6. Crenças, atitudes e motivação dos estudantes

A avaliação das crenças e das atitudes dos estudantes frente ao PI, bem como de fatores relacionados a sua motivação, foram discutidos por uma grande quantidade de artigos encontrados em nossa revisão ($N=20$). As opiniões dos estudantes foram coletadas através de questionários (usualmente com escalas do tipo *likert*) e entrevistas. Três trabalhos utilizaram questionários específicos, a saber: *Motivated Strategies for Learning Questionnaire* – MSLQ [90]; *Course Interest Survey* – CIS [91]; *Colorado Learning Attitudes about Science Survey* – CLASS [92].

Os resultados encontrados apontam que os estudantes:

- Avaliam positivamente o feedback obtido através dos Testes Conceituais [83, 93];
- Percebem que o engajamento proporcionado pelo PI foi importante para sua aprendizagem [11, 15, 18, 19, 46];
- Aferem que o PI aumentou seu interesse em estudar Física [11, 43];
- Sentem que o PI modificou seus hábitos de estudo [11];
- Percebem que o PI motivou sua presença e participação ativa em aula [11, 15, 49, 93];
- Avaliam de maneira positiva a importância da interação com os colegas para a aprendizagem dos conceitos [10, 12, 17–19, 48, 54, 71, 93, 94];
- Aprovaram o uso do PI [18, 19, 71, 77];
- Percebem que o PI promoveu um relacionamento mais positivo entre os estudantes e entre os estudantes e o professor [54, 71, 77];
- Sentem que o PI foi responsável por sua melhora na aprendizagem dos conteúdos [15, 48, 74, 77, 94];
- Aprovaram o uso do PI com outras metodologias interativas de ensino [65, 74];
- Avaliaram que a interação com os colegas aumentou sua confiança na resolução de problemas [76].

A lista de crenças e atitudes, apresentadas anteriormente, demonstra uma grande variedade de per-

cepções dos estudantes em termos dos impactos proporcionados pela adoção do PI. Podemos perceber que a adoção do PI desenvolve, de maneira geral, sentimentos positivos em vários aspectos relacionados com a aprendizagem dos conteúdos. Apenas o trabalho de Gok [16] não encontrou diferenças na motivação dos estudantes que tiveram aulas com o PI em comparação aos estudantes que tiveram aulas tradicionais.

3.2.7. Crenças de autoeficácia

Entre os fatores que afetam a maneira como os estudantes se engajam no processo de aprendizagem, as crenças de autoeficácia [95] vem ganhando certo destaque na pesquisa na grande área STEM [96]. O conceito de autoeficácia está relacionado ao julgamento pessoal sobre a própria capacidade de organizar e realizar certos tipos de tarefas visando alcançar um objetivo determinado [95]. Entre os trabalhos encontrados nesta revisão, destacamos as pesquisas de Gok [43] e de Miller et al. [97].

Gok [43] avaliou a influência da adoção do PI em um curso introdutório de Física, em uma universidade Turca, nas crenças de autoeficácia de estudantes em compreender os conteúdos abordados na disciplina, bem como do controle sobre esses sentimentos. Os resultados, encontrados através da análise de um questionário aplicado aos estudantes no final da disciplina, mostraram que os alunos que tiveram aulas com o PI reportaram crenças de autoeficácia altas, em termos do domínio do conteúdo abordado na disciplina. O questionário também foi aplicado a uma turma de controle. A análise estatística feita pelo autor apontou uma diferença estatisticamente significativa, com $p = 0.003$, entre a turma experimental e a turma de controle. Segundo o autor, o PI foi introduzido no curso para tornar os estudantes mais confiantes, ter controle sobre sua habilidade em aprender conceitos de Física e obter sucesso na aprendizagem.

Miller et al. [97], em um estudo conduzido em uma turma disciplina introdutória de Física que aborda tópicos de eletromagnetismo, analisaram a relação entre a troca de respostas dos Testes Conceituais e as crenças de autoeficácia apresentadas pelos estudantes no início do semestre. Os autores concluíram que os estudantes que possuem níveis baixos de autoeficácia tendem a modificar mais suas respostas aos Testes Conceituais, especialmente entre respostas corretas para incorretas, ou mantêm suas respostas incorretas. Esse comportamento também foi observado nos Testes Conceituais classificados como difíceis. Como sugestões, os autores indicam a

necessidade de criar intervenções que modifiquem as crenças de autoeficácia dos estudantes para que estes tenham uma melhor experiência com o PI.

3.2.8. Metacognição

Com o objetivo de avaliar os impactos da adoção do PI e do uso de sistemas de votação (sem a metodologia PI) na metacognição dos estudantes, Jones et al. [49] conduziram um estudo em uma turma de Entomologia. Através da aplicação do teste denominado “*Metacognitive Awareness Inventory*” (MAI) [98], como pré e pós-teste, os autores buscaram avaliar diferenças entre as duas propostas metodológicas na construção de dois aspectos da metacognição: conhecimento da cognição e regulação da cognição. O aumento no conhecimento da cognição foi mais expressivo nos estudantes que tiveram aulas com o PI. Já em relação à regulação da cognição, ambos os grupos tiveram um aumento similar. No entanto, ao analisar os gêneros e os tratamentos (PI e não PI), os autores encontraram que os homens do grupo que não usaram o PI tiveram um ganho significativo na regulação da cognição, enquanto no caso das mulheres, foi o grupo que usou o PI que apresentou ganho na regulação da cognição. Segundo esses autores, tal resultado é coerente com a literatura da área que afirma que mulheres preferem ambientes colaborativos de ensino [99].

3.2.9. Pesquisas referentes aos elementos que compõem o PI

O Peer Instruction pode ser considerado uma metodologia que faz uso de diversos elementos, dando forma a sua estrutura. Alguns trabalhos tiveram como foco elementos específicos do método, focando no(a): tipo de sistema de votação utilizado [100]; tipo de interação que ocorre na etapa de discussão entre os colegas [94, 101–103]; importância das atividades prévias às aulas [87]; tempo de submissão das respostas dos estudantes [104]; apresentação dos resultados da primeira rodada de votação [105, 106]; relação entre a localização dos estudantes e a distribuição de repostas [107]. Além disso, um dos trabalhos encontrados nesta revisão almejou construir uma expressão analítica que revela a correlação entre a taxa de acertos, nos Testes Conceituais, dos estudantes antes e após a discussão entre os colegas [108].

As variações nos tipos de interação que ocorrem entre os estudantes, em relação à quantidade de pontos atribuídos para acertos e erros das questões conceituais, foram investigadas por James [101] e James et al [102]. Quando os estudantes recebem

uma nota maior pelo acerto, tendem a concordar mais com os colegas. Já quando a nota é dada apenas para a participação, independente de acerto ou erro, os estudantes tendem a manter diálogos onde várias opiniões são levadas em conta, havendo uma dispersão nas respostas após a segunda rodada de votação, indicando que os estudantes votam com as suas convicções. Nessa mesma linha, Chou e Lin [94] propuseram dois mecanismos para incentivar os estudantes a interagir com seus colegas, denominados de “atribuição de interlocutor” e “responsabilização de pontuação”. O primeiro envolve a atribuição prévia, realizada pelo professor, dos grupos de interação após a primeira rodada de votação. Já o segundo propõe uma divisão no escore atribuído ao acerto/erro das questões conceituais, onde o escore máximo só é obtido quando o grupo de interação, como um todo, opta pela resposta correta após a discussão. Ambas as propostas se mostraram frutíferas para promover o engajamento dos estudantes. Já James e Willoughby [103] buscaram categorizar os tipos de interação que acontecem na etapa de discussão com os colegas. Através da análise das interações dos estudantes, os autores encontraram, em muitos casos, discussões pouco frutíferas (devido à baixa compreensão dos Testes Conceituais), levando a interpretação de que as questões utilizadas não providenciaram um feedback efetivo da aprendizagem dos estudantes.

Em relação à importância da execução das atividades prévias às aulas, o trabalho de Lucas [87] apontou que o desempenho nos Testes Conceituais em aula é superior quando os estudantes realizam as atividades prévias. Ainda em relação aos Testes Conceituais, o estudo de Monterola e Roxas [107], realizado em disciplinas de Física de nove instituições de ensino, mostrou que a localização dos estudantes que possuem um desempenho superior na disciplina afeta a performance da classe, como um todo, nos Testes Conceituais. Segundo os autores, a qualidade e a frequência com que a interação entre os estudantes ocorre é influenciada pela disposição dos estudantes. De maneira geral, a transmissão da resposta correta ocorre mais facilmente quando os estudantes classificados como “bons” estão localizados nas extremidades da sala.

Os trabalhos de Perez et al. [105] e Chien et al. [106] assinalaram que a apresentação das distribuições das respostas dos Testes Conceituais pode interferir na maneira como os estudantes modificam suas respostas. De maneira geral, os estudantes que tiveram contato com a distribuição de repostas, após a primeira rodada de votação, tenderam a modificar sua resposta da de menor frequência para a de maior

frequência. Segundo Chien et al. (Ibid.), mostrar as respostas após a primeira rodada de votação limita as direções das discussões entre os estudantes, que acabam debatendo sobre as respostas mais votadas sem levar em consideração os argumentos científicos envolvidos. Em termos dos sistemas de votação utilizados pelos estudantes, o trabalho de Lasry [100] indicou que não existem diferenças na aprendizagem conceitual dos estudantes que utilizam *clickers* ou *flashcards*.

O tempo gasto pelos estudantes para votar nos Testes Conceituais foi investigado por Miller et al. [104]. Respostas corretas são registradas em um tempo menor do que respostas incorretas, antes e após a discussão entre os colegas. Os estudantes que têm conhecimento superior de Física, bem como crenças elevadas de autoeficácia, respondem mais rapidamente em ambas as rodadas de votação. Além disso, não foram encontradas diferenças de gênero no tempo de respostas dos estudantes com o mesmo nível de conhecimento de Física.

Nessa seção buscamos descrever, de maneira detalhada, os principais impactos da adoção do PI. De maneira sintetizada, e corroborando com a análise de Vickrey et al. [21], podemos perceber que o PI tem produzido, em diversas áreas do conhecimento, resultados extremamente positivos quanto a aprendizagem conceitual dos estudantes. Tal impacto é acompanhado por um aumento nas habilidades dos estudantes em resolverem problemas. Algumas consequências do aumento na aprendizagem conceitual, bem como da habilidade de resolução de problemas, podem ser percebidas no aumento do desempenho acadêmico dos estudantes, no aumento da retenção dos estudantes nas disciplinas e nos cursos que almejam graduarem-se, e no desenvolvimento de habilidades metacognitivas. Outros resultados positivos incluem o desenvolvimento de crenças e atitudes extremamente positivas em relação ao método, um aumento na motivação dos estudantes para aprender e o desenvolvimento de maiores níveis de autoeficácia.

3.3. Impactos Instrucionais

Nessa seção, buscamos abordar os resultados de pesquisas que objetivam avaliar variações em implementações do PI, bem como investigar as atitudes dos professores frente a metodologia.

3.3.1. Variações na implementação

Nessa subseção, apresentamos pesquisas em que a estrutura do PI, proposta por Mazur [22], sofreu algum tipo de mudança, assim como as que intro-

duzem indicadores para classificar variações nas implementações do PI.

Miller et al. [79] investigaram a implementação do PI feita por 14 professores diferentes em 17 “*small classes*” de cálculo. As variações nas implementações feitas por tais instrutores foram classificadas em termos dos tipos de ConcepTests utilizados, bem como com a frequência que tais tipos de questões foram utilizados. Nessa mesma linha, Turpen e Finkelstein [109], através de um estudo etnográfico com seis professores de Física, encontraram variações na implementação do PI, as quais foram caracterizadas pelos tipos de Testes Conceituais utilizados, pelo número de Testes Conceituais utilizados por aula, pela adoção de nota pela participação durante as rodadas de votação, pela interação entre o professor e o aluno, pelo tempo dado aos estudantes para responder às questões, pelo tempo discutindo as respostas dos ConcepTests, pela participação dos alunos na discussão final e pela discussão das alternativas incorretas. Os professores observados lecionavam disciplinas introdutórias de Física, em classes de 130 a 240 estudantes, sendo dois deles professores novatos a respeito da adoção do PI e o restante membros ativos do grupo de pesquisa em ensino de Física da universidade em que foi conduzido o estudo. Em todas as aulas observadas, os estudantes testaram e aplicaram os conceitos, bem como discutiram os conteúdos com seus colegas. Contudo, diferentes oportunidades foram dadas aos estudantes para formular e responder questões, para avaliar a exatidão das respostas, para interagir com os professores e para discutir conceitos científicos com toda a classe.

Turpen e Finkelstein [110], em uma continuidade do trabalho de Turpen e Finkelstein [109], investigaram se as diferentes normas sociais construídas nas aulas de três professores, que utilizaram variações do PI, eram percebidas pelos estudantes. As diferentes normas sociais foram construídas em termos das ações dos professores e dos alunos; especificamente, do nível de interação entre os alunos e o professor, do nível de interação dos estudantes entre si e da maneira como os estudantes respondem aos testes conceituais. As diferentes maneiras como os três professores investigados no estudo fizeram dos tipos de interação, bem como da ênfase dada às justificativas dos testes conceituais, construíram três diferentes microculturas. A percepção dos estudantes frente às microculturas foi extremamente ligada aos aspectos elencados pelos pesquisadores. Em outras palavras, as diferentes maneiras como os professores se apropriaram do PI puderam ser percebidas pelos estudantes.

Artigos que investigam a inclusão do PI em novas propostas didáticas, ou seu uso concomitante com outras metodologias de ensino, que se baseiam na atividade colaborativa, também foram encontrados nesta revisão. Howard et al. [111] relataram a investigação da adoção de um jogo denominado “Survivor”, que utiliza o PI como uma de suas etapas, em uma disciplina de Fisiologia. Outros tipos de modificação incluem o uso conjunto com outras estratégias didáticas, a exemplo o trabalho de Gray e Steer [66], que reportaram o uso conjunto do PI com os “Lecture Tutorials”. Outros exemplos são os trabalhos de Cahyadi [67], Busch [65] e Gok [112], que fizeram o uso do PI em conjunto com as metodologias: *Active Learning Problems Sets*, *Constructivist Classroom Dialogues* e *Demonstrations* (Cahyadi [67]); *SCALE-UP* e *Just-in-time Teaching* (Busch [65]); *Stepwise Problem-Solving Strategie* (SPSS) (Gok[112]).

Alguns trabalhos propuseram modificações na estrutura do método. Miller [17] relatou a inclusão de discussões no grande grupo após a segunda etapa de votação, aproximando-se, dessa forma, da proposta do TPS [3]. Ainda em termos da etapa de discussão entre os colegas, James e Willoughby [103], Arteaga e Vinken [83] e James e Willoughby [103] introduziram a etapa de discussão entre os colegas já na primeira (e única) rodada de votação, ou seja, o professor apresentava o Teste Conceitual, os alunos discutiam entre eles e, então, votavam. As modificações propostas por Cox e Junkin [47] também incluíam a discussão entre os colegas para a primeira rodada de votação. Além disso, por tratar-se de aulas de laboratório de Física, os alunos foram divididos em grupos, que votavam através de um computador. Caso ocorresse divergência nas respostas, os alunos discutiam em pares de grupos.

Outra modificação na estrutura original do PI é a inclusão de uma questão conceitual após a segunda rodada de votação, que aborde o mesmo conceito que a questão anterior, porém que apresente a pergunta de forma diferente, denominada questão isomórfica [61]. Entre os trabalhos abarcados nesta revisão da literatura, as pesquisas conduzidas por Lee et al. [56] e Zingaro e Porter [113] utilizaram questões isomórficas.

3.3.2. Crenças e atitudes dos professores

As crenças e atitudes dos estudantes frente ao PI foram amplamente investigadas. No entanto, encontramos apenas uma pesquisa que reporta as crenças e atitudes dos professores que fizeram uso da me-

todologia. Rudolph et al. [11] buscaram elencar as opiniões dos professores que participaram, de forma voluntária, da implementação do PI em aulas de Física em uma universidade francesa. A receptividade do método foi extremamente positiva. Ao final do semestre, os 15 professores responderam a um questionário e participaram de uma reunião para refletir sobre suas respectivas experiências com o PI. Entre as perguntas, os professores deviam atribuir uma nota entre 1 e 5 para a eficiência do PI na aprendizagem dos alunos e para a motivação para participar das aulas; como resultado obtiveram, respectivamente, 3.8 (n=15) e 3.6 (n=14). Os comentários dos professores ressaltam que o PI proporcionou uma maior interatividade na sala de aula, onde a interação entre os estudantes, bem como dos estudantes com o professor, foi intensificada, inclusive abrindo espaço para que os alunos explicassem os conteúdos para seus colegas durante o período de aula. O *feedback* imediato da aprendizagem dos estudantes também foi um aspecto elencado pelos professores participantes desse estudo.

Em termos dos impactos instrucionais, nossa análise aponta que nem todas pesquisas analisadas em nosso trabalho utilizaram o PI na sua forma original. Alguns trabalhos apresentaram discussões acerca de modificações feitas na estrutura original do método. Algumas dessas modificações foram realizadas para investigar de maneira mais profunda a aprendizagem dos estudantes, a exemplo do uso de questões isomórficas, e para adaptar a metodologia para outros ambientes de ensino, como a adoção do PI em aulas de laboratório de Física. Outras modificações foram feitas em termos das etapas de votação e do momento onde a interação entre os colegas ocorria. Além disso, alguns trabalhos fizeram o uso do PI em conjunto com outras metodologias interativas de ensino. Tais resultados mostram que o PI é uma metodologia que pode ser facilmente adaptada, de acordo com os objetivos traçados.

Adicionalmente, os professores apresentam uma boa receptividade em relação à metodologia, especialmente pois a estrutura do PI favorece a obtenção de um *feedback* instantâneo da aprendizagem dos estudantes. Indubitavelmente, os resultados obtidos pelas pesquisas com o PI mostram inúmeros benefícios de sua adoção.

3.4. Referencial Teórico e Metodológico

Um importante resultado de nossa revisão foi encontrado através da análise dos referenciais teóricos e metodológicos adotados nas pesquisas em relação ao

PI. Nessa seção apresentamos uma discussão acerca da maneira como as pesquisas encontradas em nossa revisão foram conduzidas.

3.4.1. Referencial Teórico

Dos 72 trabalhos revisados, apenas 10 mencionaram, de maneira explícita, a adoção de algum referencial teórico para dar suporte aos achados empíricos apresentados. Um exemplo é a investigação conduzida por Gok [43], a qual averiguou as crenças de autoeficácia dos estudantes em relação a capacidade de aprender Física. O referencial adotado foi a Teoria Social Cognitiva de Bandura [95], em específico, os aspectos relacionados às crenças de autoeficácia. Os resultados encontrados mostraram que os estudantes que tiveram aulas com o PI desenvolveram crenças de autoeficácia superiores aos alunos que tiveram aulas tradicionais. Segundo o autor, o engajamento durante o processo de aprendizagem proporcionado pelo PI auxiliou os estudantes a ganharem confiança em, e ter controle sobre, suas capacidades em aprender o conteúdo da disciplina. O estudo de Miller et al. [97] também utilizou a teoria de Bandura para investigar a relação entre as crenças de autoeficácia dos estudantes e a maneira com que estes modificam suas respostas após a interação entre os colegas. Os autores concluíram que estudantes com baixos níveis de autoeficácia tendem a modificar suas respostas com maior frequência.

O Modelo de Recursos para Aprendizagem (*resources model of learning*) de Hammer [114] foi adotado como referencial teórico em duas pesquisas. Koretsky e Brooks [53] usaram tal modelo para analisar as interações discursivas que ocorrem entre os estudantes após a primeira rodada de votações dos testes conceituais. Segundo os autores, o diálogo entre os estudantes desempenha um papel crucial na ativação de recursos cognitivos apropriados, permitindo que os alunos percebam o problema de maneira diferente e, portanto, respondam aos testes conceituais corretamente. Wood et al. [115] investigaram as interações discursivas entre os estudantes na etapa de discussão entre os colegas, especificamente avaliando o processo de desenvolvimento conceitual dos estudantes através das interações com os pares. Segundo os autores (Ibid.), o Modelo de Recursos para Aprendizagem é uma ferramenta útil para identificar desalinhamentos entre os objetivos de uma questão conceitual e os efeitos que ela causa nas discussões entre os alunos.

O conceito de Normas Sociais [116] foi utilizado por Turpen e Finkelstein [110] ao investigar diferentes tipos de adoção do PI feita por professores. As

variações das normas sociais da sala de aula foram descritas dentro de três dimensões que integram o PI: colaboração entre professor e estudante; colaboração entre estudantes; construção de sentido e de respostas para os testes conceituais.

Aspectos relacionados às Redes Neurais [117] foram utilizados por Monterola e Roxas [107] para investigar a influência da distribuição dos estudantes na sala de aula no desempenho da turma em Testes Conceituais. Os autores demonstraram que uma Rede Neural é capaz de caracterizar a performance da turma em um Teste Conceitual, baseando-se nos locais onde os estudantes estão sentados. Quando os estudantes que têm um alto desempenho na disciplina estão localizados nas extremidades da sala de aula, a transmissão da resposta correta ocorre mais facilmente.

A categorização de diferentes tipos de Testes Conceituais foi feita à luz da Taxionomia dos Objetivos Educacionais de Bloom [118] no trabalho conduzido por Rao e DiCarlo [70]. Os Testes Conceituais foram classificados dentro dos dois grandes grupos de domínio cognitivo: retomada de informações e atividades intelectuais.

Algumas pesquisas, ao investigar os impactos do PI na aprendizagem dos estudantes ou, até mesmo, ao tentar dar um suporte para a eficiência do método, citam algumas teorias de aprendizagem. Suppaitayaporn et al. [119] recorrem à Teoria da Mudança Conceitual de Posner [120] para afirmar que o PI é uma metodologia que facilita o processo de mudança conceitual. A obtenção do *feedback* da aprendizagem dos estudantes, através do uso dos Testes Conceituais, foi indicada como importante frente à Teoria das Hierarquias de Aprendizagem de Gagné [121] no estudo de Jones et al. [49]. Já a eficiência da etapa de discussão entre os colegas foi ressaltada por Lenaerts et al [13], sobre ótica do conceito de Zona de Desenvolvimento Proximal de Vygotsky [122]. Apesar de citarem teorias de aprendizagem para dar suporte às suas afirmações, o delineamento da pesquisa e a análise dos resultados não as levam em conta.

3.4.2. Referencial Metodológico

Em todos os trabalhos incluídos em nossa revisão, a metodologia de pesquisa utilizada foi explicitada. Dos 72 artigos encontrados em nossa revisão, 52 utilizam análises quantitativas, envolvendo análises estatísticas de diferenças e análises descritivas, nove utilizam análises quantitativas e qualitativas e nove valem-se apenas de análises qualitativas.

O enfoque metodológico dos artigos levantados, especialmente os que tratam de investigar a eficiência do PI comparando grupos experimentais, com o uso da metodologia, e grupos de controle, com aulas expositivas é, predominantemente, quantitativo. Conforme já afirmamos, os artigos, geralmente, baseiam-se na análise estatística das diferenças entre escores em testes padronizados, bem como na análise dos ganhos normalizados dos grupos.

As investigações das opiniões de estudantes e professores, acerca da metodologia empregada, foram conduzidas, em sua totalidade, através de análises qualitativas. Alguns trabalhos utilizaram metodologias qualitativas específicas, como é o caso dos trabalhos de Turpen e Finkelstein [109, 110], que foram conduzidos sobre o enfoque de estudos de caso etnográficos, e o trabalho de Wood et al. [115], que utilizou a Teoria Fundamentada nos Dados de Strauss e Corbin [123].

Em termos dos referenciais teóricos e metodológicos adotados, nossa revisão indica que, frente à grande quantidade de trabalhos que investigaram os impactos da adoção do PI em termos da aprendizagem dos estudantes, poucos valeram-se de referenciais teóricos de aprendizagem para subsidiar os achados empíricos. Em consonância, tais trabalhos, em sua quase totalidade, adotaram como estratégia metodológica estudos experimentais, onde análises estatísticas foram utilizadas para responder às questões de pesquisa. De forma alguma desqualificamos tais trabalhos, apenas ressaltamos que investir em pesquisas que utilizem referenciais teóricos de aprendizagem, bem como metodologias de pesquisa que utilizem dados qualitativos, se faz necessário para solidificar os achados empíricos em termos dos impactos da adoção do PI.

4. Conclusões

O PI é uma metodologia de ensino que foi desenvolvida, inicialmente, para melhorar a aprendizagem conceitual dos estudantes em cursos introdutórios de Física pelo professor Eric Mazur. O sucesso obtido pelas primeiras adoções dessa metodologia, bem como a aspiração dos professores em modificarem suas práticas docentes, motivou muitos pesquisadores a aplicarem o PI em diversas disciplinas e contextos educacionais. A produção de conhecimento acerca dos impactos da adoção do PI é extremamente variada. Passados quase 25 anos de sua criação, buscamos, no presente trabalho, realizar uma meta-análise da produção acadêmica em torno do PI. A diversidade de resultados encontrados mostra que

o PI é uma proposta altamente recomendada, especialmente quando se busca engajar o estudante no processo de aprendizagem, bem como melhorar seus resultados de aprendizagem.

Em nossa revisão, podemos identificar uma grande concentração de publicações em alguns temas de pesquisa, tais como os impactos da adoção do PI na aprendizagem conceitual dos estudantes, nas habilidades de resolução de problemas e, também, nas atitudes dos estudantes frente a metodologia. A importância dos resultados apresentados nesses estudos é incontestável. Contudo, o número reduzido de estudos conduzidos em alguns contextos e temas de investigação podem ser compreendidos como *gaps* da literatura. Nesse sentido, mais pesquisas são necessárias para avaliar a adoção do PI em ambientes formais de educação básica e em disciplinas das áreas de Ciências Médicas e Ciências Humanas a nível universitário. Da mesma maneira, pesquisas nas temáticas relacionadas à retenção dos estudantes (na disciplina e no curso de graduação), ao desenvolvimento de habilidades metacognitivas nos estudantes, bem como às crenças e atitudes dos professores frente a metodologia, são necessárias para uma melhor compreensão dos impactos do PI.

Esperamos que a discussão das pesquisas, incluídas nessa revisão, contribua para que pesquisadores se motivem a investigar a adoção do PI em seus contextos de ensino. Igualmente, acreditamos que nossa síntese, através das quatro questões de pesquisa estabelecidas, forneça respostas para as principais dúvidas que um pesquisador pode ter em relação ao estado da arte da produção do PI.

Agradecimentos

Ives Solano Araújo agradece ao CNPq pela bolsa de produtividade em pesquisa concedida.

Referências

- [1] M. Prince, *Journal of Engineering Education* **93**, 223 (2004).
- [2] D.R. Stead, *Active Learning in Higher Education* **6**, 118 (2005).
- [3] F.T. Lyman, in: *Mainstreaming Digest: A Collection of Faculty and Student Papers*, editado por A.S. Anderson (University of Maryland Press, Maryland, 1981).
- [4] H.S. Barrows e R.M. Tamblyn, *Problem Based-Learning: An Approach to Medical Education* (Springer, New York, 1980), v. 1, p. 224.
- [5] L.K. Michaelsen e M. Sweet, *New Directions for Teaching and Learning* **2011**, 41 (2011).
- [6] G.M. Novak, E.T. Patterson, A.D. Gavin e W. Christian, *Just-In-Time Teaching: Blending Active Learning with Web Technology* (Prentice Hall, New Jersey, 1999), v. 1, p. 188.
- [7] R.J. Beichner, J.M. Saul, D.S. Abbott, J.J. Morse, D.L. Deardorff, R.J. Allain, S.W. Bonham, M.H. Dancy e J.S. Risley, in: *PER-Based Reform in University Physics*, editado por E.F. Redish e P.J. Cooney (American Association of Physics Teachers, Maryland, 2007).
- [8] C. Henderson, M.H. Dancy e M. Niewiadomska-Bugaj, *Physical Review Special Topics - Physics Education Research* **8**, 20104 (2012).
- [9] M. Dancy e C. Henderson, *American Journal of Physics* **78**, 1056 (2010).
- [10] C.H. Crouch e E. Mazur, *American Journal of Physics* **69**, 970 (2001).
- [11] A.L. Rudolph, B. Lamine, M. Joyce, H. Vignolles e D. Consiglio, *Physical Review Special Topics - Physics Education Research* **10**, 1 (2014).
- [12] D.J. Nicol e J.T. Boyle, *Studies in Higher Education* **28**, 457 (2003).
- [13] J. Lenaerts, W. Wieme e E. Van Zele, *European Journal of Physics* **24**, 7 (2003).
- [14] N. Lasry, E. Mazur e J. Watkins, *American Journal of Physics* **76**, 1066 (2008).
- [15] S. Butchart, T. Handfield e G. Restall, *Teaching Philosophy* **32**, 1 (2009).
- [16] T. Gok, *Asia-Pacific Forum on Science Learning and Teaching* **13**, 1 (2012).
- [17] T.K. Miller, *Issues in the Undergraduate Mathematics Preparation of School Teachers* **3**, 1 (2013).
- [18] M.G. Müller, R.V. Brandão, I.S. Araujo e E.A. Veit, *Caderno Brasileiro de Ensino de Física* **29**, 491 (2012).
- [19] V. Oliveira, E.A. Veit e I.S. Araujo, *Caderno Brasileiro de Ensino de Física* **32**, 180 (2015).
- [20] A. de Barros, J. Remold, G.S.F. da Silva e J.R. Tagliati, *Revista Brasileira de Ensino de Física* **26**, 63 (2004).
- [21] T. Vickrey, K. Rosploch, R. Rahmanian, M. Pilarz e M. Stains, *CBE-Life Sciences Education* **14**, 1 (2015).
- [22] E. Mazur, *Peer Instruction: A User's Manual* (Prentice Hall, Upper Saddle River, 1997), v. 1, p. 253.
- [23] D. Hestenes, M. Wells e G. Swackhamer, *The Physics Teacher* **30**, 141 (1992).
- [24] I.S. Araujo e E. Mazur, *Caderno Brasileiro de Ensino de Física* **30**, 362 (2013).
- [25] E. Mazur e J. Watkins, in: *Just in Time Teaching: Across the Disciplines, and Across the Academy*, editado por M. Maier e S. Simkins (Stylus Publishing, Sterling, 2009).
- [26] B. Clyde, F. Herreid, N.A. Schiller, K.F. Herreid e C.B. Wright, *Journal of College Science Teaching* **44**, 75 (2014).

- [27] E.E. Prather e G. Brissenden, *Astronomy Education Review* **8**, 10103 (2009).
- [28] J. Immerwahr, *Teaching Philosophy* **32**, 233 (2009).
- [29] D. Berrett, *The Chronicle of Higher Education* **31**, A16 (2012).
- [30] R.A. Bloodgood, *Anatomical Sciences Education* **5**, 367 (2012).
- [31] B.T. Jacques-Fricke, A. Hubert e S. Miller, *Journal of College Science Teaching* **39**, 24 (2009).
- [32] J. Szlachta, *The Journal of Nursing Education* **52**, 355 (2013).
- [33] M.F. Golde, C.L. McCreary e R. Koeske, *Journal of Chemical Education* **83**, 804 (2006).
- [34] C.M. Sorensen, A.D. Churukian, S. Maleki e D.A. Zollman, *American Journal of Physics* **74**, 1077 (2006).
- [35] B.B. Bodemer, *College & Research Libraries* **75**, 162 (2014).
- [36] A.V. Vázquez, K. McLoughlin, M. Sabbagh, A.C. Runkle, J. Simon, B.P. Coppola e S. Pazicni, *Journal of Chemical Education* **89**, 1025 (2012).
- [37] A.W.M. Lee, J.K.Y. Ng, E.Y.W. Wong, A. Tan, A.K.Y. Lau e S.F.Y. Lai, *Journal of Chemical Education* **90**, 388 (2013).
- [38] K. Woelk, *Journal of Chemical Education* **85**, 1400 (2008).
- [39] M. Wilkerson, W.G. Griswold e B. Simon, in: *Proceedings of the 36th SIGCSE technical symposium on Computer science education - SIGCSE '05*, St. Louis, 2005, editado por W. Dann, T. L. Naps, P. T. Tymann e D. Baldwin (ACM Press, New York, 2005), p. 116.
- [40] S. Ouko, C. Aurah e M. Amadalo, *Journal of Education and Practice* **6**, 175 (2015).
- [41] D.P. Maloney, T.L. O'Kuma, C.J. Hieggelke e A. Van Heuvelen, *American Journal of Physics* **69**, S12 (2001).
- [42] J.C. Libarkin e S.W. Anderson, *Journal of Geoscience Education* **53**, 394 (2005).
- [43] T. Gok, *International Journal of Science and Mathematics Education* **10**, 417 (2012).
- [44] R.R. Hake, *American Journal of Physics* **66**, 64 (1998).
- [45] S. Pilzer, *PRIMUS* **11**, 185 (2001).
- [46] L.A. Van Dijk, G.C. Van Der Berg e H. Van Keulen, *European Journal of Engineering Education* **26**, 15 (2001).
- [47] A.J. Cox e W.F. Junkin, *Physics Education* **37**, 37 (2002).
- [48] B. Schmidt, *European Journal of Engineering Education* **36**, 413 (2011).
- [49] M.E. Jones, P.D. Antonenko e C.M. Greenwood, *Journal of Computer Assisted Learning* **28**, 477 (2012).
- [50] G. Zhu e C. Singh, *Physical Review Special Topics - Physics Education Research* **8**, 1 (2012).
- [51] J. Morice, N. Michinov, M. Delaval, A. Sideridou e V. Ferrières, *Journal of Computer Assisted Learning* **31**, 722 (2015).
- [52] B.J. Brooks e M.D. Koretsky, *Journal of Chemical Education* **88**, 1477 (2011).
- [53] M.D. Koretsky e B.J. Brooks, *International Journal of Engineering Education* **27**, 897 (2011).
- [54] A. Lucas, *PRIMUS* **19**, 219 (2009).
- [55] B.J.T. Morgan e C. Wakefield, *Journal of Science Teacher Education* **41**, 51 (2012).
- [56] C.B. Lee, S. Garcia e L. Porter, *ACM Transactions on Computing Education* **13**, 1 (2013).
- [57] A.E. Relling e M.J. Giuliodori, *AJP: Advances in Physiology Education* **39**, 167 (2015).
- [58] N. Michinov, J. Morice e V. Ferrières, *Computers & Education* **91**, 1 (2015).
- [59] D. Zingaro e L. Porter, *Computers & Education* **71**, 87 (2014).
- [60] L. Porter, C.B. Lee, B. Simon e D. Zingaro, in: *Proceedings of the Ninth Annual International Conference on International Computing Education Research - ICER '11*, Providence, 2011, editado por K. Sanders (ACM, New York, 2011), p. 45.
- [61] M.K. Smith, W.B. Wood, W.K. Adams, C. Wieman, J.K. Knight, N. Guild e T.T. Su, *Science* **323**, 122 (2009).
- [62] D.E. Meltzer e K. Manivannan, *American Journal of Physics* **70**, 639 (2002).
- [63] P.S. Shaffer e L.C. McDermott, *American Journal of Physics* **60**, 994 (1992).
- [64] P.S. Shaffer e L.C. McDermott, *American Journal of Physics* **60**, 1003 (1992).
- [65] H.C. Busch, *The Physics Teacher* **48**, 578 (2010).
- [66] B.K. Gray e D.N. Steer, *Journal of College Science Teaching* **41**, 80 (2008).
- [67] V. Cahyadi, *Higher Education Research & Development* **23**, 455 (2004).
- [68] C.S. Kalman, M. Milner-Bolotin e T. Antimirova, *Canadian Journal of Physics* **88**, 325 (2010).
- [69] L.E. Keiner e T.E. Burns, *The Physics Teacher* **48**, 108 (2010).
- [70] S.P. Rao e S.E. DiCarlo, *AJP: Advances in Physiology Education* **24**, 51 (2000).
- [71] R.N. Cortright, H.L. Collins e S.E. DiCarlo, *AJP: Advances in Physiology Education* **29**, 107 (2005).
- [72] M.W. Petersen, L.N. Toksvang, R.R. Plovsing e R.M.G. Berg, *AJP: Advances in Physiology Education* **38**, 101 (2014).
- [73] J.P. Hartmann, L.N. Toksvang e R.M.G. Berg, *AJP: Advances in Physiology Education* **39**, 120 (2015).
- [74] M.C. Schuller, D.A. DaRosa e M.L. Crandall, *Academic Medicine: Journal of the Association of American Medical Colleges* **90**, 384 (2015).
- [75] G.R. Pike, *Assessment Update* **9**, 10 (1997).
- [76] T. Gok, *Journal of Baltic Science Education* **12**, 747 (2013).

- [77] M.J. Giuliadori, H.L. Lujan e S.E. DiCarlo, *AJP: Advances in Physiology Education* **30**, 168 (2006).
- [78] B.D. Hestenes e M. Wells, *The Physics Teacher* **30**, 159 (1992).
- [79] R.L. Miller, E. Santana-Vega e M.S. Terrell, *PRIMUM* **16**, 193 (2006).
- [80] D. Steer, D. McConnell, K. Gray, K. Kortz e X. Liang, *Science Educator* **18**, 30 (2009).
- [81] C.T. Richardson e B.W. O'Shea, *American Journal of Physics* **81**, 231 (2013).
- [82] S.P. Rao e S.E. DiCarlo, *AJP: Advances in Physiology Education* **25**, 127 (2001).
- [83] I.L. Arteaga e E. Vinken, *European Journal of Engineering Education* **38**, 652 (2013).
- [84] D. Zingaro, *ACM Transactions on Computing Education* **15**, 1 (2015).
- [85] K.A. Roth, *PRIMUM* **22**, 353 (2012).
- [86] K. Arnesen, G.S. Korpås, J.E. Hennissen e J.B. Stav, *Electronic Journal of e-Learning* **11**, 169 (2013).
- [87] A.R. Lucas, *PRIMUM* **22**, 97 (2012).
- [88] A.C. Strenta, R. Elliott, R. Adair, M. Matier e J. Scott, *Research in Higher Education* **35**, 513 (1994).
- [89] B.J. Watkins e E. Mazur, *Journal of College Science Teaching* **24**, 36 (2013).
- [90] P.R. Pintrich, D.A.F. Smith, T. Garcia e W. J. Mckeachie, *Educational and Psychological Measurement* **53**, 801 (1993).
- [91] C. Kim e J.M. Keller, *British Journal of Educational Technology* **39**, 36 (2008).
- [92] W.K. Adams, K.K. Perkins, N.S. Podolefsky, M. Dubson, N.D. Finkelstein e C.E. Wieman, *Physical Review Special Topics - Physics Education Research* **2**, 10101 (2006).
- [93] S. Ghosh e F. Renna, *The Journal of Economic Education* **40**, 354 (2009).
- [94] C.Y. Chou ed P.H. Lin, *British Journal of Educational Technology* **46**, 839 (2015).
- [95] A. Bandura, *Psychological Review* **84**, 191 (1977).
- [96] M. David, S. Farro e S. S. Canetto, *Analyses of Social Issues and Public Policy* **1**, 347 (2013).
- [97] K. Miller, J. Schell, A. Ho, B. Lukoff and E. Mazur, *Physical Review Special Topics - Physics Education Research* **11**, 1 (2015).
- [98] G. Schraw and R.S. Dennison, *Contemporary Educational Psychology* **19**, 460 (1994).
- [99] J.S. Brotman and F.M. Moore, *Journal of Research in Science Teaching* **45**, 971 (2008).
- [100] N. Lasry, *The Physics Teacher* **46**, 242 (2008).
- [101] M.C. James, *American Journal of Physics* **74**, 689 (2006).
- [102] M.C. James, F. Barbieri and P. Garcia, *Astronomy Education Review* **7**, 37 (2008).
- [103] M.C. James and S. Willoughby, *American Journal of Physics* **79**, 123 (2011).
- [104] K. Miller, N. Lasry, B. Lukoff, J. Schell and E. Mazur, *Physical Review Special Topics - Physics Education Research* **10**, 1 (2014).
- [105] K.E. Perez, E.A. Strauss, N. Downey, A. Galbraith, R. Jeanne and S. Cooper, *CBE Life Sciences Education* **9**, 133 (2010).
- [106] Y.T. Chien, Y.H. Lee, T.Y. Li and C.Y. Chang, *Eurasia Journal of Mathematics, Science and Technology Education* **11**, 1089 (2015).
- [107] C. Monterola and R.M. Roxas, *Complexity* **14**, 26 (2009).
- [108] H. Nitta, *Physical Review Special Topics - Physics Education Research* **6**, 20105 (2010).
- [109] C. Turpen and N.D. Finkelstein, *Physical Review Special Topics - Physics Education Research* **5**, 20101 (2009).
- [110] C. Turpen and N.D. Finkelstein, *Physical Review Special Topics - Physics Education Research* **6**, 20123 (2010).
- [111] M.G. Howard, H.L. Collins, and S.E. DiCarlo, *AJP: Advances in physiology education* **26**, 30 (2002).
- [112] T. Gok, *International Journal of Science and Mathematics Education* **13**, 561 (2015).
- [113] L. Porter and D. Zingaro, in: *Proceedings of the 45th ACM Technical Symposium on Computer Science Education - SIGCSE '14*, Atlanta 2014, edited by J.D. Dougherty and K. Nagel (ACM, New York, 2014), p. 295.
- [114] D. Hammer, *American Journal of Physics* **64**, 1316 (1996).
- [115] A.K. Wood, R.K. Galloway, J. Hardy and C.M. Sinclair, *Physical Review Special Topics - Physics Education Research* **10**, 20107 (2014).
- [116] M. Cole, *Cultural Psychology: A Once and Future Discipline?* (Harvard University Press, Cambridge, 1998), v. 1, p. 401.
- [117] R. Tadeusiewicz, *Control Engineering Practice* **3**, 746 (1995).
- [118] B.S. Bloom, *Taxonomy of Educational Objectives Book 1: Cognitive Domain* (Longman, New York, 1956), v. 1, p. 256.
- [119] D. Suppapittayaporn, N. Emarat and K. Arayathanikul, *Research in Science & Technological Education* **28**, 63 (2010).
- [120] G. Posner, K. Strike, P. Hewson and W. Gertzog, *Science Education* **66**, 211 (1982).
- [121] R.M. Gagné and M. Driscoll, *Learning and Teaching (Principles of Educational Psychology Series)* (Prentice Hall College, Upper Saddle River, 1975).
- [122] L.S. Vygotsky, *Mind in Society: The Development of Higher Psychological Processes* (Harvard University Press, New York, 1978), p. 105.
- [123] A. Strauss and J. Corbin, *Basics of Qualitative Research: Techniques and Procedures for Developing Grounded Theory* (SAGE Publications, Thousand Oaks, 2008), v. 3, p. 379.