

COMO CONTRIBUIR NO ENSINO DO ELETROMAGNETISMO A PARTIR DAS DIFICULDADES CONCEITUAIS DOS ALUNOS?

HOW TO CONTRIBUTE TO TEACHING ELECTROMAGNETISM FROM STUDENTS CONCEPTUAL DIFFICULTIES?

Marcel Bruno Pereira Braga¹

 ORCID iD: [0000-0003-2226-1538](https://orcid.org/0000-0003-2226-1538)

Igor Rodrigues Chicolet da Silva²

 ORCID iD: [0000-0001-8076-9108](https://orcid.org/0000-0001-8076-9108)

RESUMO

Como melhorar o ensino do Eletromagnetismo? Essa é uma pergunta que pode ser respondida sob diversas perspectivas, entretanto, o objetivo do trabalho visa responder a partir de um diagnóstico sobre as dificuldades conceituais dos alunos de ensino médio e ingressantes universitários. O diagnóstico é construído a partir das evidências obtidas na administração de um teste conceitual denominado Electricity and Magnetism Concept Assesment (EMCA). Os procedimentos de análise baseiam-se na Teoria Clássica dos Testes (TCT), incluindo Fator de Hake para verificar impactos entre grupos não correspondentes. Os grupos se caracterizam e se distinguem entre perfis de proficiência e dificuldades conceituais de alunos da rede pública na cidade de Manaus. Seguindo procedimentos similares aos autores, os resultados corroboram que mesmo no nível superior há uma baixa compreensão de conceitos básicos no Eletromagnetismo. Partindo disso, comparam-se os níveis de proficiência entre os diferentes grupos normativos algumas instruções pedagógicas seguem na perspectiva de David Ausubel, visando contribuir para o desenvolvimento de aprendizagem significativa tendo como referência os itens com boa discriminação e com maiores dificuldades. Informações mais detalhadas para o grupo normativo são obtidas através da Análise Gráfica do Item (AGI), sendo destacados os distratores mais atrativos, sinalizando evidências de modelos mentais incorretos.

Palavras-chave: Eletromagnetismo. Dificuldades. Ensino.

ABSTRACT

How to improve the teaching of Electromagnetism? This is a question that can be answered from different perspectives, however, the objective of this paper is to answer from a diagnosis about the conceptual difficulties of high school students and university freshmen. The diagnosis is based on the evidence obtained from the administration of a conceptual test called Electricity and Magnetism Concept Assesment (EMCA). The analysis procedures are based on the Classical Test Theory (TCT), including Hake Factor to verify impacts between unmatched groups. The groups are characterized and distinguished between proficiency profiles and conceptual difficulties of public school students in the city of Manaus. Following similar procedures to the authors, the results corroborate that even at the

¹ Doutor em Educação Científica e Tecnológica pela UFSC. Professor Adjunto do Departamento de Física da Universidade Federal do Amazonas – UFAM, Manaus, Amazonas, Brasil. Av. General Rodrigo Octavio Jordão Ramos, 1200 – Coroado I, Manaus – AM, 69067-005. E-mail: marcelbraga@ufam.edu.br.

² Especialista em Mídias da Educação pela UFAM. Mestrando do Programa de Pós-Graduação em Ensino de Ciência e Matemática da Universidade Federal do Amazonas – UFAM, Manaus, Amazonas, Brasil. Rua Pico das Águas, 457 – Nossa Senhora das Graças, Manaus – AM, 69053-093. E-mail: chicolet.fisica@gmail.com.

higher level there is a poor understanding of basic concepts in electromagnetism. Based on this, the levels of proficiency between the different normative groups are compared. Some pedagogical instructions follow David Ausubel's perspective, aiming to contribute to the development of meaningful learning with reference to items with good discrimination and with greater difficulties. More detailed information for the normative group is obtained through the Graphical Item Analysis (AGI), highlighting the most attractive distractors, signaling evidence of incorrect mental models.

Keywords: Eletromagnetism. Difficulty. Teaching.

1 INTRODUÇÃO

Não há ensino sem aprendizagem. Porém, há uma ilusão de se ensinar Física que se mantêm muitas vezes sem a percepção e consciência de sua finalidade, geralmente pela falta de uma reflexão sobre as influências de sua prática diante o desempenho dos alunos, ou pela ingenuidade de se isentar desse processo. Essa ilusão leva a possíveis distorções na prática docente que se difundem, como por exemplo, a forte crença de professores quanto a suficiência no uso de métodos decorativos ou algorítmicos na resolução de problemas como processo a garantir que ocorra um avanço relevante na aprendizagem, geralmente sendo explorado indiscriminadamente em qualquer etapa, sendo uma ingenuidade agravante diante das dificuldades na aquisição de conceitos e no desenvolvimento de competências científicas. Dentre outras práticas ineficientes comuns estão os métodos verificativos de leis e princípios que ocorrem dentro de laboratórios experimentais de modo inapropriado, seja explorando a construção de gráficos para constatar relações entre grandezas discrepância quanto aos erros, dando pouca ênfase na interpretação física do fenômeno. Possivelmente isso ocorra por influência em um ensino diretivo condicionado a ser orientado por manuais e roteiros rígidos que acabam por instruir e efetivar uma prática descontextualizada, com procedimentos mecanizados, baixo poder de criticidade, discussão e reflexão. Algumas dessas distorções são condutas que configuram crenças ingênuas docentes enraizadas a caracterizarem uma abordagem acrítica, aproblemática e descontextualizada, entre outros, que distanciam de uma visão coerente de ciência, seja por ação ou omissão (Gil-Pérez et al., 2005).

Um dos maiores estudos sobre educação baseado em evidências (Hattie, 2017) aponta que uma das influências mais diretas a impactar negativamente o desempenho dos alunos, esteja relacionada a uma postura docente em ignorar as suas dificuldades e necessidades, pode-se afirmar com isso que este pode ser um “divisor de águas” para o alcance dos objetivos educacionais preestabelecidos, sendo considerado um ponto crucial e uma variável diretamente relacionada ao fracasso. No campo prático-experimental do ensino, entende-se que essas variáveis se relacionam na insistente tentativa de explorar uma compreensão física

demasiadamente na abstração, dissociando-a da observação e investigação fenomenológica. Além disso, ressalta-se sobre a importância de uma reflexão prévia diante o conhecimento a ser envolvido, sobre o ensino e aprendizagem em curso, bem como, após o processo e de perspectivas futuras.

Aderindo a uma perspectiva consoante ao *design educacional*, entende-se que haja uma necessidade de realizar um levantamento e utilização adequada na elaboração de modelos de ensino a aprendizagem a partir das percepções dos alunos, pois tornam-se subsídios norteadores imprescindíveis para um planejamento e o exercício de um *ensino sob medida*, com previsões de êxito dos estudantes e a predição das dificuldades na compreensão de conteúdos específicos e habilidades cognitivas associadas, entre outros elementos que auxiliam na tomada de decisão que possibilite uma melhoria do ensino. Nesse paradigma, torna-se difícil esperar bons resultados na aprendizagem, sem que seja valorizado as necessidades legítimas cognitivas dos alunos como elementos a conduzir práticas formativas psicologicamente coerentes. Para isso, a aquisição de conceitos básicos e estruturadores tornam-se “ingredientes” na preparação de um “terreno fértil” para uma aprendizagem em níveis mais abstratos, configurando assim como parâmetros estruturadores e articuladores para a compreensão de um conjunto mais vasto e complexo do conhecimento, exercendo assim um papel pedagógico estratégico para o desenvolvimento de competências científicas mais complexas.

Quanto as implicações instrucionais para o ensino a partir da valorização na análise dessas dificuldades e desempenho, Ausubel (2003) sugere que a aquisição de conceitos seja significativa partindo preferencialmente de um corpo de conhecimento mais inclusivo e estruturado, na qual os conceitos básicos devem se relacionar de forma estável e cada vez mais complexa na estrutura cognitiva. A organização e o desenvolvimento das etapas de aprendizagem, segundo Ausubel, deve contemplar em especial os processos de *diferenciação progressiva e reconciliação integradora*. Para isso, os níveis de compreensão desses conceitos devem se relacionar de forma lógica entre os diversos conceitos subjacentes e situações, sem perder de vista a forma e o contexto em que eles se mobilizam diante das leis e princípios mais gerais, e se de fato é um *contexto significativo para o aprendiz*. O distanciamento do contexto da realidade dos alunos tende a tornar-se também esse processo sem sentido. Vale ressaltar que a aprendizagem tende a ser significativa quando se identifica a preexistência de subsunçores (conceitos-âncoras) na estrutura cognitiva e se busca desenvolver tarefas psicologicamente significativas, utilização de materiais logicamente significativos relacionados com conteúdos ou conhecimentos significativos. Em última perspectiva, esses são os objetivos que se idealiza através do bom uso do diagnóstico advindo da pesquisa. Nesse sentido, o uso dos *testes*

conceituais remete e se apresenta como um potencial instrumento a articular essas demandas, seja na avaliação da eficiência do ensino, quanto na verificação da existência de subsunções com subsídios necessários na tomada de decisões.

Para tentar contribuir com o ensino da Física no campo do eletromagnetismo, a pesquisa assume uma preocupação em analisar o desempenho de um teste como evidências de possíveis subsunções, e com isso sinalizar as necessidades dos alunos nesse campo do conhecimento. Dentre as evidências, buscou-se destacar as informações mais consistentes e que remetem a modelos mentais equivocados e dificuldades na compreensão, com isso, o olhar se desvia dos acertos ou no alcance de domínio mínimo dentro de um conjunto de conhecimentos, apesar de que ambos estarem intrinsecamente relacionados. Nisso, considerou-se apropriado a utilização do *Electricity and Magnetism Concept Assessment – EMCA* (McCOLGAN et al, 2017) como instrumento a recolher as informações para as análises, considerando sobretudo que esteja à serviço de uma melhoria na qualidade de ensino. Para delimitar e direcionar o estudo, busca-se responder a seguinte questão: *Quais são os conteúdos associados as dificuldades conceituais em um teste conceitual, na perspectiva de David Ausubel, que possibilitam melhorar o ensino de Eletromagnetismo?*

O objetivo principal está em identificar os conteúdos potencialmente significativos a serem desenvolvidos pelo professor, a partir das análises, que possam contribuir no enfrentamento e superação das dificuldades conceituais em Eletromagnetismo.

As ações remetem para a análise das dificuldades conceituais evidenciadas no desempenho do teste, fundindo índices psicométricos em interpretações pedagógicas relevantes que possibilitem propor intervenções metodológicas mais eficientes para o ensino. Para o alcance do objetivo geral, tem-se os seguintes objetivos específicos são:

- *Identificar os subsunções relacionados as concepções alternativas nos distratores em destaque na avaliação de desempenho em cada nível de escolaridade e na comparação entre eles;*
- *Interpretar pedagogicamente as informações psicométricas mais relevantes e consistentes dos resultados;*
- *Sugerir a abordagem de conteúdos específicos do eletromagnetismo, com base nas evidências, tendo a finalidade de instruir e propor melhorias para o ensino de eletromagnetismo.*

2 REFERENCIAL TEÓRICO

2.1 O Instrumento.

O *inventário de conceitos* (“teste conceitual”) utilizado foi desenvolvido em 2010 por Michele McColgan, Rose Broder e George Hassel a qual aplicaram em na turma de Álgebra da Universidade de George Washington. Esse teste baseou-se em outros testes como o *Force Concept Inventory – FCI* (HESTENES; WELLS; SWACKHAMER, 1992) e *Brief Electricity and Magnetism Assessment – BEMA* (DING et al. 2006) para a elaboração.

Inicialmente, o teste conceitual a ser administrado seria o BEMA. No entanto, uma análise no conteúdo do teste sugeriu um nível elevado de dificuldade, e optou-se pela mudança para o EMCA, na qual encontram-se disponíveis no site <https://www.physport.org>, associado ao *American Association of Physics Teachers*.

De acordo com McColgan et al. (2017), o EMCA iniciou seu desenvolvimento em 2010 e em sua primeira versão tinha 46 itens, que foram reduzidos para uma quantidade de 30 itens de múltipla escolha. Teve como base outros *testes conceituais* como o próprio BEMA, destacando, ainda, a grande quantidade de respostas por item e o alto grau de dificuldade dos demais testes. Rahmawat et al. (2018) afirma que uma das desvantagens do BEMA é a ausência de itens a respeito de eletrodinâmica e que o teste possui uma abordagem universitária.

Segue abaixo uma tabela mostrando o inventário de conceitos abordados no EMCA:

Quadro 1 – Inventário de Conceitos do EMCA.

Conteúdo	Itens/alternativas
Condutores e Distribuição de Carga	1c, 20e, 25d
Carga	1c, 4e, 6b, 9b, 20e
Lei de Coulomb	2e, 3d, 9b
Lei de Gauss e Fluxo	5c
Campo Elétrico	6b, 7a, 8b, 25d
Força devido ao Campo Elétrico	7a
Energia Potencial Elétrica e Trabalho	9b, 10a
Circuito em série	11a, 12c, 13d
Circuito em Paralelo	14c, 15b, 16a
Circuito (resistência)	17b
Resistência	18b, 19a
Corrente induzindo um Campo Magnético	21d, 28b, 29e
Força de uma carga em movimento em um Campo Magnético	22c, 25d, 30c
Imãs	23d
Corrente induzida – Lei de Lenz	24c
Separação de carga – potencial	25d
Força em um fio	26a
Fluxo Magnético	27d
Campo Magnético	28b, 29e

Fonte: Traduzido a partir de McColgan et al. (2017).

Quanto as evidências de validade, é importante destacar a necessidade de avançar em diversos aspectos além da adequação semântica advinda da tradução, na qual se apresenta disponível em língua inglesa. Para isso, concebem-se os critérios apontados por Braga (2018), em que deve diferenciar um teste *traduzido* de um *adaptado transculturalmente*. É preciso ressaltar que não há como garantir que uma tradução literal dos itens evite a formação de frases incompreensíveis ou com interpretações semânticas distintas, ou seja, de que o traço latente seja o mesmo nas duas versões sem que seja rigorosamente realizadas as etapas de adaptação transcultural. Apesar disso, no site <https://www.physport.org> considera-se o teste com “validação bronze”, ou seja, baseado em pesquisas sobre o pensamento de estudantes, utilizando análises estatísticas apropriadas, conduzidas em múltiplas instituições e com publicação revisada por pares.

As perguntas de múltipla escolha sobre a EMCA foram desenvolvidas por um grupo de cinco membros do corpo docente com base em sua experiência no ensino dos tópicos relevantes. A EMCA foi concedida a mais de 200 estudantes de física introdutória do segundo semestre, com base em álgebra e cálculo, e análises estatísticas apropriadas de confiabilidade, dificuldade e discriminação foram conduzidas. Valores razoáveis foram encontrados. As pontuações para homens e mulheres foram semelhantes. (ver em: <https://www.physport.org/assessments/assessment.cfm?I=89&A=EMCA>. Tradução livre).

2.2 Alfa de Cronbach

Conforme Eaton et al. (2019), que fez aplicações utilizando o BEMA e CSEM, diz que a confiabilidade do teste refere-se à consistência que este teste tem em dar uma nota específica para o respondente; ressalta que valores abaixo de 0,8 fazem com que o teste conceitual esteja apto a relacionar grupos, mas não indivíduos. O Alfa de Cronbach é igual a:

$$\alpha = \left(\frac{n}{n-1} \right) \cdot \left(1 - \frac{\sum_{i=1}^n \sigma_i^2}{\sigma_t^2} \right)$$

onde, n → número de itens; σ_i → variância do item e σ_t → variância total do teste.

3 METODOLOGIA

A pesquisa se caracteriza como uma investigação exploratória tendo como base um processo de *avaliação de desempenhos*, na qual busca mapear dificuldades conceituais em contextos diferenciados de ensino e na proposição de instruções pedagógicas sob medida para o ensino do Eletromagnetismo. Essas dificuldades e proposições educacionais seguem sob a

perspectiva teórica da psicologia educacional de David Ausubel. A pesquisa decorreu no município de Manaus entre os meses de maio a junho de 2019. Foram incluídas duas escolas estaduais da rede pública do Ensino Médio (EM), tendo como amostra no *pré-teste* ($n_{pré}=92$) alunos do 1º ($n_1=37$), 2º ($n_2=22$) e 3º ano ($n_3=33$), bem como discentes da disciplina de Física no Ensino Superior (ES) da Universidade Federal do Amazonas (UFAM) como parte da amostra do *pós-teste* ($n_{pós}=35$). Os respondentes do ES se caracterizam em níveis e cursos diversificados.

Diante dos dados, foram utilizadas análises clássicas dentro da Psicometria para a avaliação dos desempenhos nos testes, nomeadamente, os procedimentos de análises contemplam a Teoria Clássica dos Testes, sendo complementadas com a *Análise Gráfica do Item* (VAN BATENBURG; LAROS, 2002) e o *Fator de Hake* (HAKE, 1998).

Devido ao curto tempo para desenvolver a pesquisa, as observações foram emparelhadas e transversais, ou seja, com uma única aplicação dos testes. As implicações diante disso resultou em um *design* na qual o grupo de pré-teste foram os alunos do ensino médio e o pós-teste, os alunos universitários, ou seja, de grupos não correspondentes. O estudo também recolheu o consentimento livre e esclarecido dos participantes.

4 ANÁLISES E RESULTADOS

Para os dois grupos em que o teste foi aplicado, o do pré-teste obteve um Alfa de Cronbach igual a 0,25, e o do pós-teste de 0,60. Apesar de haver uma diferença significativa entre os índices, ambos são considerados índices baixos, aceitável para análises e algumas medidas projetivas (MURPHY; DAVIDSHOFER, 2005, p. 150), porém mais adequado para avaliar e comparar médias de grupos. O valor entre 0,41 e 0,60 também é considerado como moderado (LANDIS; KOCH, 1977). Entre os fatores que influenciam sobre a classificação de consistência está o procedimento de aplicação, o comprometimento dos respondentes e uma possível validade interna duvidosa do instrumento. Nisso, tem-se o número de questões, a redundância entre itens, e a correlação entre os itens elevam o número do alfa. No entanto, no que tange ao nível de pesquisa proposta, os valores são considerados aceitáveis e suficientes, uma vez que se busca estimar perspectivas diante de médias de grupos não correspondentes.

4.1 Zona de Guessing (Acerto Aleatório)

Para determinar a “guessing rate” do teste (Braga, 2018) deve-se fazer uma breve separação dos itens quanto ao número de alternativas, pois cada grupo de itens terá uma probabilidade de acerto ao acaso diferente. Em seguida, multiplica-se pela quantidade de itens que fazem parte daquele grupo. Com isso, o “guessing rate” do teste usado neste projeto, o EMCA, que possui 30 itens, sendo dois itens com 3 alternativas, oito com 4 e o restante com 5 alternativas ($k=5$), é de 7 itens ($2 \cdot 0,3 + 8 \cdot 0,25 + 20 \cdot 0,2 = 6,67 \sim 7 \text{ itens}$), correspondendo a 23,4% do teste, e entende-se que o **desempenho mínimo esperado** seria a *metade do caminho* para o alcance máximo além do acerto aleatório do teste, ou seja: $7 + \frac{30-7}{2} = 18,5 \sim 19 \text{ itens}$, correspondendo a 63,4% do teste, sendo essa uma linha de corte para a zona de domínio conceitual sobre os fundamentos de eletromagnetismo, ao contrário de 50% como geralmente se considera. O “guessing do item” (BRAGA, 2018) é o mesmo para um conjunto de respondentes do item, o que corrobora com a ideia de que a taxa de adivinhação do item é a mesma independentemente da quantidade de respondentes, no caso, se associa com a perspectiva da probabilidade de acerto aleatório (A_a) do item com K alternativas: $P(A_a) = \frac{1}{k}$. Ressalta-se que, para respondentes que habitam a zona de “guessing do item”, é pouco fiável afirmar que estes dominam o conhecimento avaliado, apesar de não se deve generalizar. De modo semelhante, admite-se que o *alcance percentual mínimo esperado* para um item seria a metade do caminho para o alcance máximo além da probabilidade de acerto aleatório do item, no caso, para $K = 5$, seria: $\frac{1}{5} + \frac{4}{2} = 0,20 + 0,40 = 0,60$ (60%).

Tabela 1 – Guessing dos itens

Itens	“Guessing do item” ou Probabilidade do acerto aleatório do item $[P(A_a)]=1/k$	Probabilidade do erro $[P(e)]=1-1/k$	Alcance percentual mínimo esperado
y1, y9	0,33	0,67	0,67
y10, y11, y12, y14, y15, y18, y19, y24	0,25	0,75	0,63
y2, y3, y4, y5, y6, y7, y8, y13, y16, y17, y20, y21, y22, y23, y25, y26, y27, y28, y29, y30	0,20	0,80	0,60

Fonte: Elaborado pelos autores.

Pode-se observar (Gráfico 1) de modo geral ($n=175$) que o percentual médio de acerto dos itens estão aquém da linha de corte para um domínio mínimo esperado (em torno de 60%), estando em sua maioria abaixo das linhas de corte do *guessing* do item. Lembrando que essa amostra contempla alunos do 1º, 2º e 3º com uma perspectiva diagnóstica (grupo Pré-teste,

sendo que os alunos do 3º ano já haviam assistido algumas aulas de eletromagnetismo), enquanto o alunos do ES correspondem a 20% da amostra (grupo Pós-teste não-correspondente).

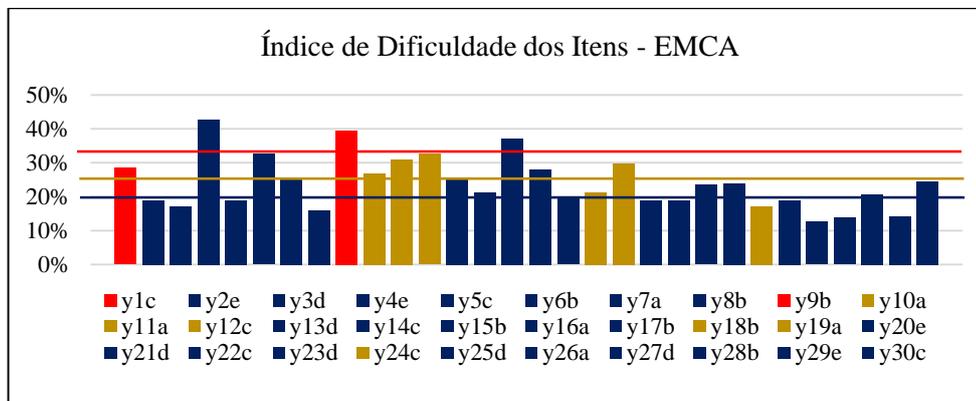


Gráfico 1 – Percentual de acerto da amostra geral (respondentes da EB e ES).

Fonte: Elaborado pelos autores.

É conveniente que se analise os itens na qual se obtiveram as maiores porcentagens acima do guessing do item, nessa perspectiva as análises cabem aos itens y9, y4, y6, y7, y9, y10, y11, y12, y15, y16, y19, y22, y23 e y30. Alguns deles encontram-se no limiar do guessing e com isso são itens pouco fiáveis, porém a qualidade do item deve ser caracterizada pelo parâmetro de discriminação. De acordo com o inventário de conceitos do teste (Quadro 1) tem-se os seguintes conteúdos abordados para os itens com maior consistência: *Leis de Coulomb* (y2e); *Carga elétrica* (y4e); *Carga Elétrica/Potência e Trabalho* (y9b); *Circuito Paralelo* (y16a); e *Resistência* (y19a)

O que percebe-se de forma significativa é que os conceitos trabalhados nos itens que se destacaram estão situados no início de todo o *Eletromagnetismo*, sugerindo que: (1) a parte do conteúdo de magnetismo não chega ao alunos; (2) parte do conteúdo de magnetismo chega com qualidade duvidosa; ou (3) o nível de abstração exigido pelo conteúdo de magnetismo não está sendo atendido pelo grupo, fazendo com que a compreensão torne-se mais dificultosa.

4.2 Índice de Dificuldade

O *Índice de Dificuldade* (ID) é, muitas das vezes, controverso, uma vez que esse parâmetro mede o percentual de acerto de um determinado item, em outras palavras, quanto maior for o ID, mais fácil será o item (BRAGA, 2018). A fim de melhorar a discriminação, propõe-se uma tabela para que as análises clássicas dos itens sejam úteis (BRAGA, 2018 apud

CERDÁ). Os IDs para o teste conceitual sugerem um teste predominantemente difícil para os respondentes (Tabela 2).

Tabela 2 – Classificação para o Índice de Dificuldade

Tipologia do item	Proporção esperada	Limite aproximado	Amostra geral (n=175)	Pré-teste (EB: n=167)	Pós-teste (ES: n=35)
Muito fáceis	10%	de 0,75 a 0,95	-	-	y4
Fáceis	20%	de 0,55 a 0,74	-	-	-
Normais/ Medianos	40%	de 0,45 a 0,54	-	-	y9, y10, y12, y16, y19
Difíceis	20%	de 0,25 a 0,44	y1, y4, y6, y7, y9, y10, y11, y12, y13, y15, y16, y19, y30	y1, y4, y6, y9, y11, y12, y15, y19	y1, y2, y3, y6, y7, y11, y13, y14, y15, y18, y20, y21, y22, y23, y28, y30
Muito difíceis	10%	de 0,05 a 0,24	y2, y3, y5, y8, y14, y17, y18, y20, y21, y22, y23, y24, y25, y26, y27, y28, y29	y2, y3, y5, y7, y8, y10, y13, y14, y16, y17, y18, y20, y21, y22, y23, y24, y25, y26, y27, y28, y29, y30	y5, y8, y17, y24, y25, y26, y27, y29

Legenda sobre o guessing dos itens: 0,33 (vermelho); 0,25 (marrom); 0,20 (azul).

Fonte: Adaptado de Braga (2018).

De modo geral, todos os 30 itens encontram-se nas zonas de tipo “difícil” e “muito difícil”, uma vez que a maior parcela dos respondentes se encontram em pré-teste (80% da amostra) e não tiveram – ou tiveram de forma superficial – o contato com o conteúdo, com isso, é natural que haja uma maior dificuldade para esse grupo normativo.

Analisando de forma isolada cada grupo normativo, pode-se perceber que, no grupo correspondente ao pós-teste, houve uma mudança razoável, mas não representativa, sendo seis itens que avançaram para uma classificação de dificuldade do tipo “normal” (y9, y10, y12, y16, y19) e “muito fácil” (y4). A mudança entre grupo era de se esperar, visto que o grupo com os universitários – que tinham a porcentagem de conclusão do curso diversificada – apresentam o conhecimento que os alunos do Ensino Básico ainda não tiveram. No entanto, vale ressaltar que deveria haver uma distribuição mais homogênea dos itens quanto às suas tipologias; ainda há um número significativo de itens “difíceis”, demonstrando que a compreensão dos conceitos abordados está longe da realidade do educando.

4.3 Índice de Discriminação

O *Índice de Discriminação* (DISCR) – também conhecido como discriminação clássica para que não seja confundido com a discriminação da *Teoria de Resposta ao Item* – é um parâmetro comparativo entre alunos de “alto” e “baixo” desempenho (EATON et al, 2019).

Para isso, deve-se organizar o desempenho dos respondentes em um rol como padrão normativo, em seguida usa-se um percentil (27% no caso) que represente os escores superiores e, o mesmo percentil, para os respondentes com escores inferiores. Considera-se determinante para a qualidade do item que ele possui uma *boa discriminação*, sendo esse o critério predominante para a seleção de itens a serem analisados. Dentro desses grupos, os respondentes com o mesmo escore devem fazer parte dessa seleção com isso, a amostra acaba não correspondendo precisamente aos 27% dentro do rol. O *DISCR* é determinado pela diferença desses parâmetros, apesar de haverem outros índices percentuais de discriminação, como o *método abreviado dos 27%* (BRAGA, 2018), na qual divide o *DISCR* pela metade de respondentes contidos no grupo superior e inferior, podendo ser considerado um ajuste de precisão. Com a combinação entre os dois índices psicométricos, torna-se possível identificar itens com informações mais consistentes diante da dificuldade (zona sombreada):

Tabela 3 – Combinação entre ID e DISCR – Amostra Geral

DISCR \ ID	Muito fáceis (de 0,75 a 0,95)	Fáceis (de 0,55 a 0,74)	Normais (de 0,45 a 0,54)	Difíceis (de 0,25 a 0,44)	Muito difíceis (de 0,05 a 0,24)
Muito bom $DISCR \geq 0,4$	-	-	-	y4	-
Bom (sujeito a aprimoramento) $0,3 \leq DISCR < 0,4$	-	-	-	y16, y19	-
Marginal (sujeito a reelaboração) $0,2 \leq DISCR < 0,3$	-	-	-	y9	y2
Inadequado (sujeito a rejeição) $DISCR < 0,2$	-	-	-	y1, y6, y7, y10, y11, y12, y13, y15, y30	y3, y5, y8, y14, y17, y18, y20, y21, y22, y23, y24, y25, y26, y27, y28, y29

Fonte: Adaptado de Braga (2018).

Predominantemente, observa-se 25 itens (83,3% do teste) inadequados e com elevada dificuldade, revelando que de modo geral os respondentes não se diferenciam diante dessas questões, ainda que possuam desempenhos distintos quanto ao teste. No entanto, é importante destacar que parte dos respondentes são de âmbito universitário e seus respectivos escores foram superiores aos demais. Situação parecida abordada (Madsen et al., 2015) que diz sobre estudantes de Física possuírem escores mais elevados devido ao desenvolvimento do pensamento físico durante o Ensino Básico.

Entre os possíveis fatores externos ao domínio conceitual a influenciar o desempenho, estão aqueles relacionados ao *design* de aplicação e processo de administração do instrumento,

como uma possível falta de rigidez na aplicação quanto as orientações prévias e gestão do tempo, falta de comprometimento no preenchimento das respostas, ineficiência da abordagem do conteúdo por parte do docente, entre outros.

4.4 Análise Gráfica do Item

A *Análise Gráfica do Item* (AGI) é uma técnica aplicada na *Curva de Resposta dos Itens* (CRI) que demonstra utilidade para análises complementares da *Teoria Clássica dos Testes* sobre a qualidade dos itens sem a necessidade do uso de *softwares* complexos (BRAGA, 2018, p. 177), como por exemplo, para contribuir na reestruturação dos testes remetendo na eliminação ou melhorias nos itens, além constatar eventual poder atrativo de distratores para um grupo de respondentes de alta performance, tornando assim, objeto promissor a ser investigado.

Considerando os itens mais consistentes (y2, y4, y9, y16 e y19), seguem abaixo algumas análises gráficas para a amostra geral (n=175). A curva que representa a alternativa correta de cada item está caracterizada na cor verde e o *distrator* mais forte, de vermelho. Destaca-se que os itens serão comentados, um a um, no próximo subtópico, deixando para este momento a análise gráfica.

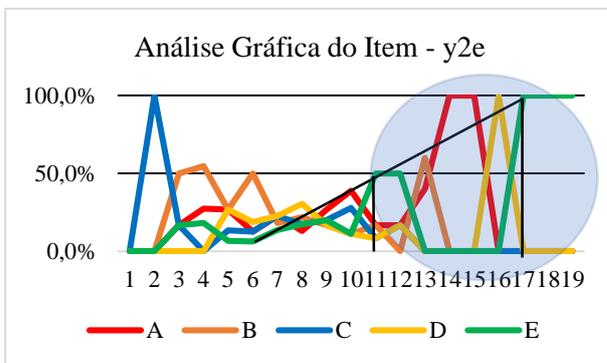


Gráfico 2 – AGI y2e na CRI ($n_{y2e}=171$).

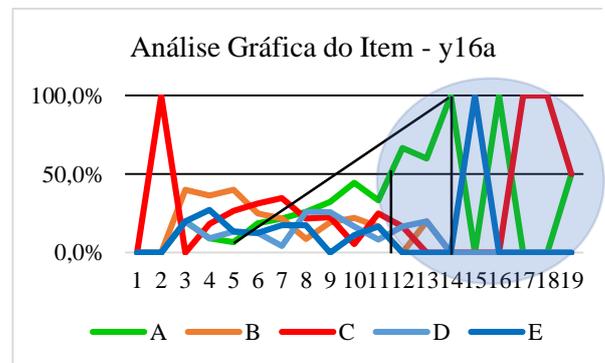


Gráfico 1 – AGI y16a na CRI ($n_{y16a}=172$).

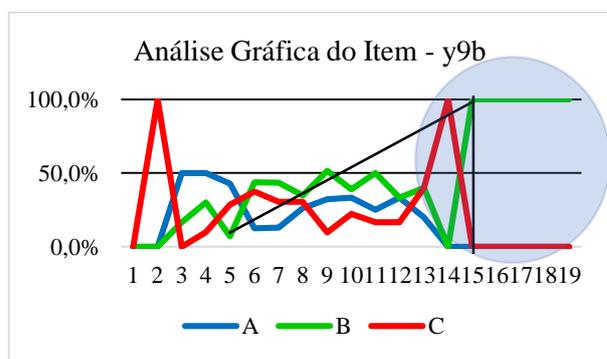


Gráfico 2 – AGI y9b na CRI ($n_{y9b}=171$).

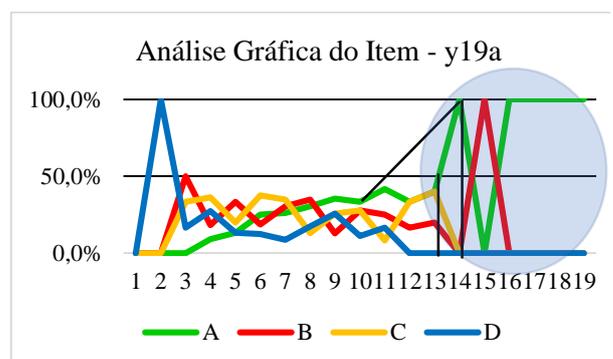


Gráfico 3 – AGI y19a na CRI ($n_{y19a}=173$).

Fonte: Elaborada pelos autores.

As análises no item y2e revelam no percentual bruto dois *distratores em destaque*: a alternativa “a” e “b”, entretanto, no AGI percebe-se que essa atração é predominante para respondentes de baixa performance, em especial para “b”, enquanto a alternativa “d” se torna atrativa para respondentes com 16 acertos, apesar de que no eescore 11, na qual o grupo alcança 50% de acertos, já se tem cerca de 96% do total de respondentes, revelando a dificuldade do item para a maioria dos respondentes.

O y9b apresenta nível de dificuldade significativo e, ainda, vê-se que um mesmo *distrator* possui dois picos de influência – tanto para respondentes com baixo eescore quanto para alto; representando uma possível má formulação do item, condizendo com a Tabela 6 sobre a reelaboração deste. O item y16, possui alta discriminação entre respondentes de alto e baixo escores. Porém, no que tange ao seu nível, é considerado difícil; percebe-se pela inclinação da linha tracejada. O que diferencia este item do próximo, y19, vai além da qualidade, é o *distrator*; onde um mesmo *distrator*, na situação a alternativa “c”, atrai a atenção dos respondentes de escores marginais.

Para respondentes de baixo escore o *distrator* representado pela alternativa “d” se destaca, enquanto para respondentes de performance mais elevada o *distrator* torna-se a alternativa “c”.

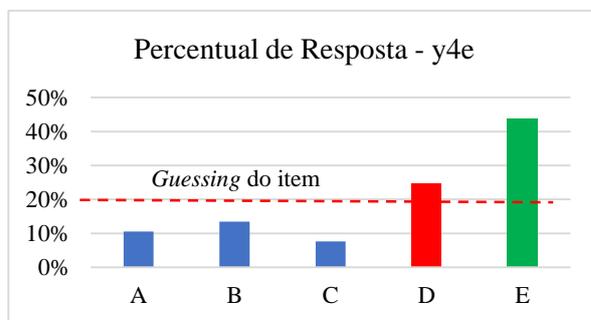


Gráfico 6 – Percentual de Resposta ($n_{y4e}=171$).

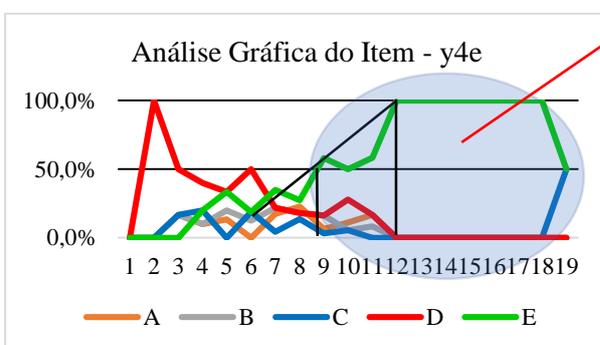


Gráfico 7 – AGI y4e na CRI ($n_{y4e}=171$).

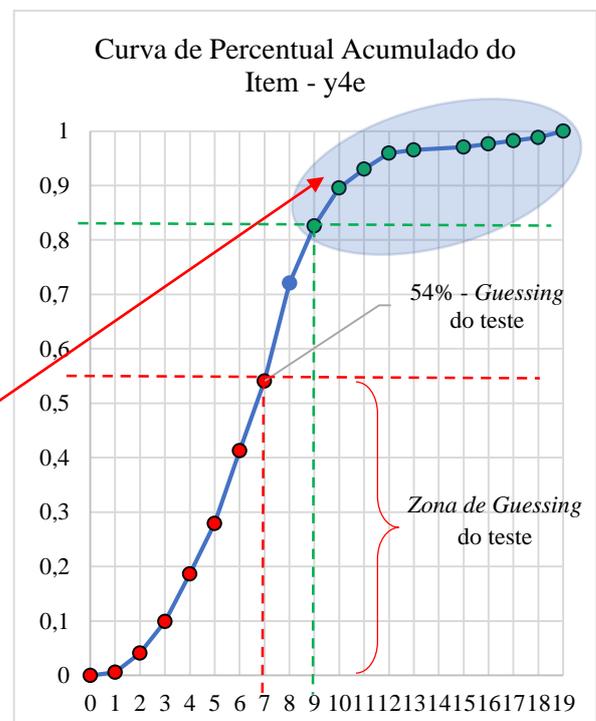


Gráfico 8 – Percentual acumulado ($n_{y4e}=171$).

Fonte: Elaborado pelos autores.

Dos cinco itens propostos à análise gráfico, o y4e é o que apresenta a maior inclinação da linha tracejada, representando uma menor dificuldade comparativamente, entretanto é classificado como “difícil”. A AGI mostra que a partir do escore 9 (83% dos respondentes) mais de 50% dos respondentes se mantem com um percentual superior de acerto, e do escore 12 em diante há uma perspectiva de predomínio quanto ao domínio conceitual avaliado. O distrator “c” possui um baixo percentual de acerto, mas chama atenção de um respondente com maior performance (escore 19) ser atraído, exatamente no limiar de domínio conceitual sobre o eletromagnetismo, que corresponde 63,4% de acertos no teste.

4.5 Ganho Normalizado

Comparando o impacto entre cursos com métodos tradicionais e ativos, Richard Hake (HAKE, 1998) apresentou evidências alunos em nível universitário apresentavam dificuldades na compreensão algumas das ideias mais fundamentais da Física, onde pode ser definido, em livre tradução, como uma *medida aproximada da eficácia de um curso na promoção da compreensão conceitual*. Sendo expresso como:

$$\langle g \rangle = \frac{\langle G \rangle}{\langle G_{m\acute{a}x} \rangle} = \frac{\langle p\acute{o}s \rangle - \langle pr\acute{e} \rangle}{100 - \langle pr\acute{e} \rangle}$$

onde, $\langle g \rangle \rightarrow$ o ganho normalizado médio; $\langle G_{m\acute{a}x} \rangle \rightarrow$ o máximo ganho possível médio; $\langle p\acute{o}s \rangle \rightarrow$ a média do pós-teste; $\langle pr\acute{e} \rangle \rightarrow$ a média do pré-teste. Apesar disso, é comum a atualização da *média dos ganhos* $g = \langle \frac{G}{G_{m\acute{a}x}} \rangle$, bem como o *tamanho do efeito* $d = \frac{\langle p\acute{o}s \rangle - \langle pr\acute{e} \rangle}{\sigma}$.

A discriminação dos impactos de ganho percentual normalizado entre os grupos de respondentes, visam entre outras avaliar a eficácia de métodos de ensino, sobretudo apresentando subsídios para a melhorias no processo pedagógico. Entende-se que este processo não exclui a possibilidade de verificar os impactos por grupos de respondentes de cada item, ao invés do teste como um todo. Conforme Hake (1998, p. 3) os impactos podem ser classificados como: $\langle g \rangle \geq 0,70$ (alto impacto); $0,70 > \langle g \rangle \geq 0,30$ (médio impacto); $\langle g \rangle < 0,30$ (baixo impacto). O gráfico sugerido por Hake para que as análises sejam feitas é o ganho normalizado $\langle g \rangle$ versus porcentagem do valor médio no pré-teste $\langle pr\acute{e} \rangle$, na qual considera que haja altos impactos já com um ganho de $\langle g \rangle \geq 0,60$. Paralelamente pode-se verificar que não diferenças substanciais entre a média dos ganhos g e o ganho médio ou ganho normalizado $\langle g \rangle$ (Fator de Hake). Considerando que o EMCA possa ser mais pertinente e adequado aos respondentes do 3º ano, percebe-se que não também diferenças comparando com os domínios de alunos do 1º e 2º ano.

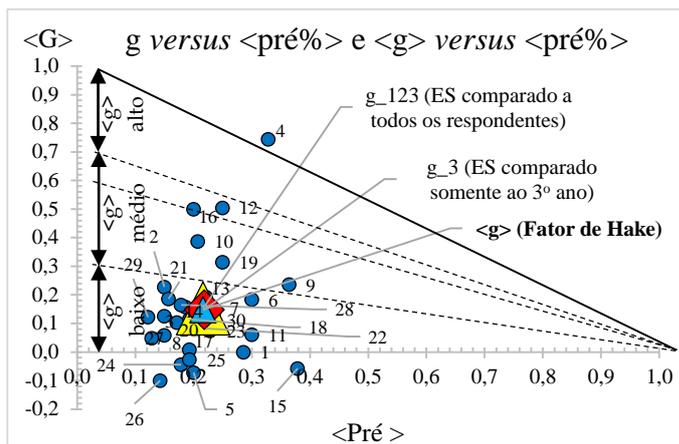


Gráfico 9 – Média dos ganhos, ganho normalizado *versus* porcentagem do valor médio no pré-teste
Fonte: Elaborado pelos autores.

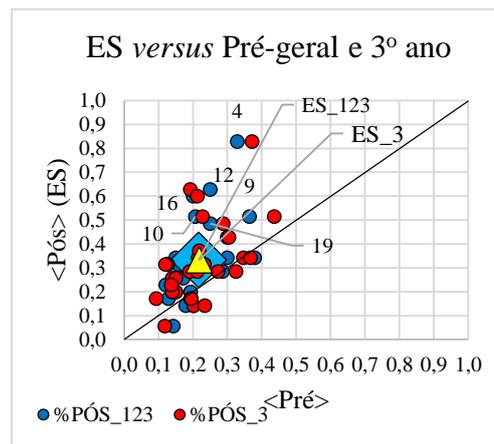


Gráfico 10 – Porcentagem pós-teste *versus* porcentagem pré-teste
Fonte: Elaborado pelos autores.

A princípio pode-se afirmar que **há um baixo impacto** comparando o domínio de estudantes universitários com os alunos de ensino médio nas escolas públicas. Percebe-se, de todos os itens, dois grupos significativos: os com médio e alto *ganho normalizado* (y4, y9, y12, y16, y19) e os que ficaram abaixo de zero, como se pudesse “desaprender” (!?) um conceito (y5, y15, y24, y25, y26), remetendo preocupação quanto a evidências da ineficiência dos processos de ensino envolvidos, entre outras influências a serem investigadas ($d = 0,03$). Os itens em verde são os de alto *ganho normalizado* e os de vermelho os que possuíram um *ganho normalizado* negativo. Os conceitos mais básicos presentes no conteúdo de *Eletromagnetismo* foram mais bem compreendidos pelos respondentes, no entanto, percebe-se que, conforme o conteúdo vai se tornando mais complexo, menos familiar os respondentes vão ficando com o conteúdo. De forma linear, a proficiência tende a ficar em torno de conhecimentos de Eletricidade, enquanto o conteúdo mais próximo do Magnetismo exige um grau de abstração que não se pode encontrar nos respondentes.

Vale destacar os itens y15 e y16, ambos destacados em negrito no quadro acima, por possuírem os mesmos conceitos envolvidos. No entanto, o item y16 tem uma abordagem mais prática e o y15 mais conceitual, sugerindo que os métodos de ensino na qual os respondentes estão submetidos são insuficientes e/ou ineficientes. O ensino mecanizado dando ênfase a uma aprendizagem baseada em modelos repetitivos e algorítmicos na resolução de problemas remete por vezes a uma ineficiência na compreensão de conceitos, do tipo “sei como resolver, mas não sei bem o que está acontecendo”, isso significa que avaliamos mais capacidades cognitivas matemáticas do que aquelas relacionadas a compreensão física, estando de lado uma preocupação sobre o entendimento dos fenômenos, dos aspectos e propriedades físicas

relacionadas as causas e efeitos. A ensino mecanizado favorece um aprendizado mecanizado, mais dedutivo que indutivo, sendo esse um modelo de ensino bem-sucedidos em exames de larga escala e vestibulares tradicionais, na qual assumem sem mérito um *status* de excelência pela eficiência em resolver problemas algorítmicos, tornando métodos e processos de ensino investigativos, interativos e integrativos, potencialmente favoráveis a uma compreensão mais profunda e consistente, um processo enfadonho e ineficiente. Nesse sentido, entra-se de forma crítica no complexo e controverso campo da avaliação educacional, que foge do escopo deste trabalho, entretanto, vale a pena refletir sobre a seguinte questão: *Qual é o ideal educativo que desejamos para o processo formativo?*

4.6 Procedimentos de Aplicação

Sobre as finalidades na utilização de *inventário de conceitos*, entende-se que possa ser explorado com diversas implicações pedagógicas, porém recomenda-se evitar como avaliações punitivas, lição de casa, ou como motivação para debates (MADSEN et. al, 2017). Sua principal finalidade esta em proporcionar uma avaliação da aprendizagem dos alunos, visando sobretudo uma eficácia para o ensino. Apesar de não medir precisamente a habilidade de resolver problemas, entende-se que certificar um domínio conceitual de base seja o mais essencial. De modo geral, defende-se boas práticas para uma melhor administração nas aplicações, algumas seriam (IBID, pp. 4-6): (1) nomear o teste com um título genérico; (2) dar o tempo recomendado para resolução; (3) promover o comprometimento da resolução, podendo atribuir uma pontuação simbólica a fim de que o motive; (4) ressaltar que o teste tem objetivo de avaliar o docente, bem como o currículo utilizado, não o indivíduo; (5) certificar de que não haja qualquer tipo de material que possa ser utilizado como consulta.

5 CONSIDERAÇÕES

Entende-se que um dos caminhos mais promissores para contribuir no ensino do eletromagnetismo está em considerar como *auxiliares* (e não obstáculos!) nas propostas de ensino as dificuldades conceituais básicas e necessárias, na qual comprometem o desenvolvimento da aprendizagem em um campo mais estruturado e complexo. Diante disso, observam-se evidências relacionadas a *discriminação (DISCR)* e *dificuldade (ID)*, na qual remetem a qualidade, consistência e capacidade avaliativa de distinguir grupos de diferentes níveis de proficiência. No caso, os valores obtidos nesse estudo são, em geral, bem discrepantes aos dados apresentados por McColgan et al. (2017), no entanto, quanto ao *DISCR*,

conhecimento -se que os valores próximos referentes aos itens y4, y9 e y16, justamente os itens com o poder de diferenciar respondentes com níveis de proficiência diferentes para um determinado conceito. Para o *ID*, os itens y1, y7, y15 e y25 apresentaram-se semelhantes, comparativamente. Partindo dessa perspectiva, e ao desconsiderar problemas relacionados a elaboração do item, os resultados sugerem que o conhecimento científico necessário para a compreensão desses itens é aparentemente independente quanto as metodologias utilizadas, pois demonstram maiores dificuldades independente do contexto, evidenciando que os conceitos envolvidos sejam talvez mais abstratos, ou que o ensino deva ser mais eficiente.

Dentro das concepções ausubelianas (MOREIRA,1999, p.162), sabemos que o professor deve exercer um papel fundamental para favorecer uma aprendizagem significativa. Com isso, torna-se relevante que o professor leve em consideração a estrutura conceitual de eletromagnetismo de modo a organizá-la de forma hierárquica, ou seja, dos conceitos menos inclusivos para os mais complexos. Nisso, os resultados obtidos se caracterizam como *subsunçores* (conhecimento prévio) relevantes a serem considerados nessa organização, de modo que as propostas de atividades possam ser psicologicamente adequadas e potencialmente significativas, mais precisamente circuitos com lâmpadas, presentes nos itens y11, y12, y13, y14, y15 e y16, semelhante as dificuldades conceituais identificadas por Eric Mazur, em Harvard (Mazur, 1997).

Essa situação sugere uma atividade prática relacionada com a elaboração concreta de um circuito formado por lâmpadas. Para isso, partir em questões abertas relacionadas aos *subsunçores* destacados torna-se um caminho promissor, podendo ser relacionadas a demonstrações investigativas e experimentação laboratorial. É importante destacar a necessidade das relações condicionantes de compreensão entre os conceitos, que muitas vezes se encontram de forma implícita. Ressalta-se que há inúmeras possibilidades de propostas pedagógicas que se ajustam a perspectiva de *David Ausubel*, como o uso de mapas conceituais e organizadores, porém a o exercício adequado de suas funções deve levar em consideração as evidências sobre o conhecimento prévio. De modo geral, entende-se que a perspectiva ausubeliana possibilita uma maior consistência na compreensão das dificuldades e de métodos mais ajustados psicologicamente. Os resultados também contribuem para uma reestruturação na projeção do teste relacionando os subsunçores dentro dos *inventários de conceitos e equívocos típicos* correspondentes no teste, remetendo ao professor vieses educacionais promissores para o desenvolvimento de atividades criativas e críticas, sobretudo, potenciais para uma aprendizagem significativa.

AGRADECIMENTOS

A todos os colaboradores das escolas pelo apoio no desenvolvimento da pesquisa. Aos alunos do PIBID/Física/UFAM (2018-2019), pela contribuição com a recolha de dados e compartilhamento de informações no grupo de estudo e pesquisa Observatório de Dificuldades na Aprendizagem Científica e Tecnológica (ODACT). E por fim, ao PIBIC/CNPq e ao PPG-ECIM.

REFERÊNCIAS

BRAGA, M. B. P. **Escala de Proficiências em Concepções Térmicas: Diagnóstico Psicométrico de Estudantes em Portugal e Brasil**. 2018. 976 f. Tese (Doutorado em Ensino das Ciências), Universidade Federal de Santa Catarina, Florianópolis, SC.

CACHAPUZ, A. et al. **A necessária renovação no ensino das ciências**. São Paulo: Cortez, 2005.

DING, L.; CHABAY, R.; SHERWOOD, B.; BEICHNER, R. **Evaluating an electricity and magnetism assessment tool**: Brief electricity and magnetism assessment, Phys. Rev. ST Phys. Educ. Res. 2 (1), 7, 2006.

EATON, P. et al. **Classical test theory and item response theory comparison of the brief electricity and magnetism assessment and the conceptual survey of electricity and magnetism**. Physical Review Physics Education Research, [s. l.], v. 15, n. 1, p. 010102, 2019. Disponível em: <https://journals.aps.org/prper/abstract/10.1103/PhysRevPhysEducRes.15.010102>. Acesso em: 15 jul. 2019.

HESTENES, D; WELLS, M; SWACKHAMER, G. **Force concept inventory**, Phys. Teach. 30 (3), 141, 1992.

HAKE, R. **Interactive-Engagement vs. Traditional Methods: A Six-Thousand-Student Survey of Mechanics Test Data for Introductory Physics Courses**. American Journal of Physics, [s. l.], 1998.

LANDIS, J. Richard; KOCH, Gary G. **The measurement of observer agreement for categorical data**. biometrics, p. 159-174, 1977.

MADSEN, A.; MCKAGAN, S. B.; SAYRE, E. C. **How physics instruction impacts students' beliefs about learning physics: A meta-analysis of 24 studies**. Physical Review Special Topics - Physics Education Research, [s. l.], v. 11, n. 1, p. 010115, 2015. Disponível em: <https://journals.aps.org/prper/abstract/10.1103/PhysRevSTPER.11.010115>. Acesso em: 16 jul. 2019.

MADSEN, A.; MCKAGAN, S. B.; SAYRE, E. C. **Best Practices for Administering Concept Inventories**. The Physics Teacher, [s. l.], v. 55, n. 9, p. 530–536, 2017. Disponível em: <https://aapt.scitation.org/doi/10.1119/1.5011826>. Acesso em: 16 jul. 2019.

MURPHY, K. R.; DAVIDSHOFER, C. O. **Psychological Testing: Principles and Applications**. Pearson Education, 2005.

MCCOLGAN, M. W. et al. **Assessing students' conceptual knowledge of electricity and magnetism**. *Physical Review Physics Education Research*, [s. l.], v. 13, n. 2, p. 020121, 2017. Disponível em:
<<https://journals.aps.org/prper/abstract/10.1103/PhysRevPhysEducRes.13.020121>>. Acesso em: 13 fev. 2019.

MOREIRA, M. A. A teoria da aprendizagem significativa de Rogers. In:_____. **Teorias da Aprendizagem**. São Paulo: EPU, 1999. cap. 9, p. 139-149. Disponível em:
<<http://leticiafrancomartins.pbworks.com/w/file/attach/97972008/Cap%209%20Moreira.pdf>>
Acesso em: 30 jul. 2019.

VAN BATENBURG, T. A.; LAROS, J. A. **Graphical analysis of test items**. *Educational Research and Evaluation*, v. 8, n. 3, p. 319-333, 2002.

Submetido em: 09 de agosto de 2019.

Aprovado em: 20 de dezembro de 2019.