

Identificação de Critérios para Avaliação do Pensamento Computacional Aplicado

César França¹, Cleudimar Galindo²

¹Departamento de Computação/Universidade Federal Rural de Pernambuco (UFRPE)

²Mestrado Profissional em Engenharia de Software, Cesar School

franssa@ufrpe.br, cgs@cesar.school

Abstract. *Computational Thinking (CT) is a problem-solving paradigm based on computer science techniques, composed of cognitive skills which are non-observable, non-precise and subjective, which represents a difficulty to evidence its learning. In this article, we report a qualitative synthesis conducted with the survey data collected from experienced computational thinking professors, which resulted in an initial set of CT assessment criteria. Our results are already useful to guide CT teachers to improve their assessment job, but can also serve as a basis for the development, in the future, of more robust and objective learning assessment models for Computational Thinking.*

Resumo. *O Pensamento computacional (PC) é um paradigma de resolução de problemas baseado em técnicas da ciência da computação, composto por habilidades cognitivas não observáveis, subjetivas e imprecisas, o que representa uma barreira para evidenciar o seu aprendizado. Neste artigo, reportamos uma síntese qualitativa de dados coletados com professores universitários experientes no ensino de PC, a qual resultou em um conjunto explícito de critérios de avaliação para cada uma das dimensões do PC. Este resultado já é útil para orientar professores, e pode também servir como base para o desenvolvimento, no futuro, de modelos de avaliação de aprendizado robustos e objetivos.*

1. Introdução

O termo Pensamento Computacional (PC) foi usado pela primeira vez por Seymour Papert, embora a professora Jeannette Wing tenha popularizado a ideia ao defender a computação pensando em todos os estudantes, como uma habilidade cognitiva básica, aplicada à resolução de problemas em qualquer área do conhecimento e para toda a vida [Wing, 2006]. Esta habilidade foi reconhecida por instituições relevantes como uma das habilidades necessárias para o profissional do futuro [DAVIS et al 2011], e desde então o termo ganhou momento na academia, e diversas pesquisas tem sido realizadas sobre como trabalhar o desenvolvimento do pensamento computacional em cursos específicos [Selby, 2014; de Araujo 2016; Avila, 2017].

Embora uma grande quantidade de pesquisas venha buscando esclarecer o conceito, e desenvolver práticas de ensino e métodos de aprendizagem, a avaliação do desenvolvimento de PC continua desafiando pesquisadores, uma vez que o PC engloba um conjunto de habilidades cognitivas sobre as quais não se há um consenso definitivo,

e ainda são de difícil aferição devido fundamentalmente à subjetividade dos conceitos envolvidos [Avila, 2017].

Buscando contribuir para a prática da avaliação de habilidades relacionadas ao pensamento computacional, nesta pesquisa, realizamos uma síntese qualitativa de dados coletados com profissionais de computação e professores universitários experientes no ensino de PC, resultando em um conjunto explícito de critérios de avaliação da aplicação de cinco dimensões do PC: pensamento algorítmico, decomposição, generalização, abstração e avaliação. Este conjunto de critérios pode representar um primeiro passo no sentido de desenvolver mecanismos que sejam capazes de atribuir uma maior objetividade aos instrumentos de avaliação de tais habilidades de PC.

O restante deste artigo está organizado da seguinte maneira: na Seção 2, apresentamos uma breve revisão da literatura focada na verificação de aprendizagem de habilidades relacionadas a pensamento computacional, ressaltando as estratégias que têm sido reportadas na literatura e as dificuldades encontradas no processo de avaliação. Na Seção 3, detalhamos os métodos de coleta e análise de dados projetados e executados nesta pesquisa. Na seção 4, apresentamos outros detalhes sobre a execução da pesquisa, bem como os critérios de avaliação resultantes do processo de síntese dos dados coletados. Por fim, na Seção 5, apresentamos discussões, conclusões e sugestões para continuidade deste trabalho em pesquisas futuras.

2. Revisão da Literatura

2.1. Pensamento Computacional (PC): Definições e Dimensões

Em síntese, o PC é um processo fundamentalmente cognitivo, que envolve raciocínio lógico pelo qual os artefatos são analisados, problemas são resolvidos, procedimentos são criados, e sistemas são compreendidos [WING, 2006]. O PC se utiliza de ferramentas e conceitos mentais da ciência da computação que ajudam as pessoas resolverem problemas, projetarem sistemas, entenderem o comportamento humano, e envolverem computadores para auxiliar na automação de uma ampla variedade de processos [NATIONAL RESEARCH COUCIL, 2010]. Como consequência, o PC pode ser aplicado a diversas áreas do conhecimento e da vida prática [Csizmadia et al., 2015], e por isso também, pesquisas apontam o crescimento em relevância do tema [DAVIS et al 2011].

Ribeiro et. al [2017] traça um paralelo para distinguir o pensamento computacional do pensamento puramente lógico, em termos de entradas e saídas. Segundo eles, o pensamento lógico é o processo de raciocínio que é aplicado sobre um conjunto de premissas, para determinar se estas premissas são verdadeiras ou falsas. Já o pensamento computacional é uma generalização desse processo, que é capaz de receber qualquer entrada, e descrever o processo através do qual esta entrada se transformaria em uma determinada saída, de maneira clara e não ambígua.

Davis et. al [2011] define o PC como “a habilidade de traduzir grandes quantidades de dados em conceitos abstratos e compreender o raciocínio baseado em dados”. Porém, PC não se refere ao aprendizado específico de conceitos de desenvolvimento de algoritmos computacionais, mas trata-se de um conjunto de habilidades cognitivas. Diversos trabalhos encontrados na literatura ainda tratam essas duas coisas de maneira similar, como é o exemplo de Seiter & Foreman [2013], Román-

González [2015] e Rodriguez [2017]. Moreno-León [2016], por exemplo, demonstra experimentalmente que avaliações utilizadas para mensurar pensamento computacional indiretamente através da prática da escrita de código são convergentes com as mesmas métricas utilizadas para avaliar complexidade de código genericamente.

No entanto, é importante fazer a distinção, uma vez que já há na literatura, uma abundância de técnicas e maneiras de verificar o aprendizado de algoritmos concretos. Ribeiro et. al [2017] apresenta uma estrutura que define PC de maneira composta, organizada em três dimensões de habilidades de pensamento: análise, abstração e automação. Csizmadia et al. [2015], por sua vez, identifica cinco habilidades: (1) pensamento algorítmico, (2) decomposição, (3) generalização, (4) abstração e (5) avaliação. Embora existam outros trabalhos que buscam identificar outras dimensões de pensamento computacional, este modelo de cinco habilidades é o mais largamente reconhecido, e por isso foi utilizado como principal referência no desenvolvimento da presente pesquisa [SELBY e WOOLARD, 2013].

Muito embora exista o argumento de que ao se apropriar da prática da elaboração de algoritmos concretos, os estudantes desenvolvem as habilidades de PC por consequência (especialmente o pensamento algorítmico), os objetos avaliativos em questão são distintos, e os modelos de verificação precisa de aprendizagem destas habilidades cognitivas do PC é que são escassos na literatura. Na seção seguinte apresentamos uma breve revisão sobre tais modelos.

2.2. Verificação de Aprendizagem e Pensamento Computacional

Para Valente [2016], as pesquisas relativas ao pensamento computacional encontradas na literatura podem ser divididas em praticamente três grandes blocos: a natureza do pensamento computacional e como ele pode ser avaliado (como identificar o pensamento computacional no aprendiz); a formação de educadores para desenvolverem atividades que exploram os conceitos do pensamento computacional, especialmente integrados as atividades curriculares; as ideias sobre o pensamento computacional podem ser trabalhadas em conjunto com as disciplinas do currículo. Neste artigo, estamos interessados expressamente no primeiro grupo.

Brennan e Resnil [2012] desenvolveram um framework para avaliação da aprendizagem em PC, baseado em conceitos, práticas e perspectivas. Para eles, os conceitos centrais de PC envolvem: sequências, loops, paralelismo, condições, e operadores, enquanto as práticas incluem teste e depuração, reuso e modularização. Eles revisitam, então, diferentes formas de avaliar esses conceitos e práticas com base fundamentalmente na expressão da habilidade de programação. No entanto, como discutido na seção anterior, este é um exemplo de trabalho que não resolve diretamente o problema que estamos interessados.

Buscando resolver esse problema, Selby [2014] propôs um mapeamento entre disciplinas concretas do currículo de computação e o seu consequente desenvolvimento de cinco habilidades do PC. Porém, esse framework tem claras limitações no que diz respeito ao seu uso em uma disciplina focada no desenvolvimento do PC, porque pressupõe que o pensamento computacional é na verdade resultante de um currículo mais duradouro e complexo que envolve disciplinas tais como Algoritmos, Programação, Hardware, Redes e comunicação e Tecnologia da Informação.

De Araujo et. al [2016] realizou então um estudo de mapeamento da literatura, tendo revisado 51 artigos cuidadosamente selecionados a partir de um universo de quase 500 relatórios de pesquisa publicados em bases científicas, que incluíam os termos “*computational Thinking*” e “*assess OR analyze OR evaluation OR measure OR validation*”. Como resultado, também foi evidenciado que cursos de programação são a abordagem mais comum para promover o desenvolvimento do PC. Como consequência, o código é o artefato mais comumente avaliado. Também foi evidenciado que a análise do conteúdo de exercícios de resolução de problemas pode ser uma abordagem efetiva, porém não existem formas consolidadas para mensuração de tais habilidades cognitivas, por causa da natureza fundamentalmente subjetiva dos conceitos envolvidos.

Avila [2017] realizou outra revisão sistemática para investigar especificamente abordagens de avaliação utilizadas em disciplinas de PC, e evidenciou em 58 estudos que a maioria dos trabalhos reportados na literatura realizam avaliações com base em instrumentos confeccionados pelos próprios pesquisadores, sendo que diversos estudos não reportam de forma explícita qualquer fundamento teórico dos seus processos avaliativos como um todo, reforçando a ideia de que avaliação em PC permanece como um desafio, e que ainda é necessário identificar-se ou desenvolver-se métodos mais genéricos para mensurar com precisão a efetividade de do aprendizado de conceitos, habilidades e competências relacionadas ao PC.

3. Métodos

Como mencionado anteriormente, o objetivo geral desta pesquisa foi identificar um conjunto de critérios pragmáticos que possam ser utilizados para evidenciar as habilidades de pensamento computacional (PC), e conseqüentemente possa ajudar a melhor fundamentar os processos de avaliação. Para atingir este objetivo, desenhamos um procedimento de pesquisa descritiva-classificatória, para identificar os padrões de critérios de avaliação utilizados por um conjunto de professores universitários experientes no ensino de PC.

Inicialmente, realizamos um *survey* para coleta de respostas para um conjunto de questões utilizadas em provas de PC, retiradas site www.computacional.com.br. (Tabela 1). Foi elaborado um formulário online na plataforma Google Forms®, e acrescentamos questões para mapeamento do nível de formação, profissão e idade dos participantes. Realizamos uma coleta de dados piloto, para avaliar aspectos relacionados à entendibilidade das questões, e usabilidade do formulário, e finalmente o questionário foi distribuído a um conjunto de profissionais da área de computação, por meio de *whatsapp* e listas de emails.

As respostas coletadas a partir deste formulário foram então utilizadas na elaboração de um segundo *survey*, encaminhado para um conjunto de cinco professores universitários, associados a universidades públicas de Pernambuco e Paraíba, pré-selecionados por conveniência, a partir dos contatos dos autores desta pesquisa. Os cinco professores foram selecionados por já terem ministrado previamente disciplinas de PC, ou por possuírem pesquisas publicadas sobre o assunto. Dois destes professores possuíam ainda experiência com desenvolvimento de PC no ensino médio. O propósito desta atividade foi então fazer com que cada professor elaborasse um parecer sobre as provas de um subconjunto dos profissionais respondentes. Cada professor recebeu cinco provas completas.

Tabela 1 - Questões Utilizadas

Cód. - Questão
Q1: Você precisa adquirir um carro, quais os critérios que você utiliza para a compra?
Q2: Quando você precisa atar um cadarço de um tênis, como você faz? Descreva os passos.
Q3: Você está indo pescar, como você se prepara, quais as ferramentas e passos que você pretende utilizar para obter êxito?
Q4: No cenário que se mostra [ILUSTRAÇÃO] temos: 2 camisas e uma bermuda brancos, 3 camisas vermelhas e duas toalhas de banho azuis. Como você faria para lavar estas roupas em uma máquina de lavar?
Q5: Você irá realizar uma viagem de férias, como você planeja a viagem?
Q6: Descreva quais os passos que você faria para plantar uma árvore.

Tabela 2 - Dimensões do PC

Pensamento algorítmico (PA): refere-se à capacidade de especificar uma solução em termos de passos claros e bem definidos, com regras que possam ser seguidas precisamente, para resolver o problema.
Abstração (AB): diz respeito à capacidade de simplificação dos problemas, ignorando detalhes irrelevantes para a solução, e escolhendo os aspectos certos para resolver os problemas da maneira mais fácil.
Avaliação (AV): é o processo de garantir que uma solução é adequada para um determinado propósito, observando aspectos como correteude, desempenho, facilidade de uso, entre outras.
Decomposição (DE): diz respeito à habilidade de enxergar e decompor as partes de um problema, para que estas possam ser compreendidas, e resolvidas em separado.
Generalização (GE): está associada à capacidade de identificar padrões, similaridades e conexões e explorar essas características, transferindo soluções entre contextos e criando soluções genéricas.

Os avaliadores foram então orientados a verificar se as respostas dadas pelos alunos, em cada questão, atendiam ou não às cinco habilidades de PC através de uma escala simplificada: (0) não é possível avaliar, (1) não atende, (2) atende parcialmente, (3) atende totalmente.. Utilizamos como base as cinco habilidades identificadas por Selby [2014], e definidas por Csizmadia et. al [2015] (Tabela 2). Em seguida, foram incentivados a descrever uma justificativa para as suas avaliações. Por fim, realizamos um procedimento de síntese temática a partir destas justificativas, para identificar os critérios que eles levaram em conta para fundamentar as suas interpretações.

De acordo com Cruzes e Dyba [2014], a síntese temática é uma abordagem que é frequentemente utilizada na identificação de padrões, análise e projeção de padrões (temas) a partir de dados de pesquisa qualitativa primária. Para realização da síntese temática, Cruzes e Dyba [2014] recomendam ainda o seguinte procedimento: leitura inicial dos dados/textos, identificação de segmentos relevantes, rotulação dos segmentos de texto, comparação entre rótulos e códigos para redução de sobreposições, generalização em temas e criação de modelos de alta ordem a partir dos temas identificados. Esta pesquisa seguiu este conjunto de passos para realização da síntese temática.

4. Resultados e Discussões

O *survey* para coleta de respostas orgânicas ao questionário-prova de PC foi divulgado, e recebeu respostas, durante o período de 19 de Abril a 07 de Maio de 2020. Foram coletadas um total de 22 respostas, todas advindas de profissionais formados, ou em formação, em cursos da área de computação, com idades variando de 21 a 45 anos ($\mu = 34,0$, $dp = 6,4$). Dentre as profissões, autodeclaradas pelos participantes, tivemos Engenheiros de Software ($n = 5$), Analistas de Sistemas ($n = 4$), Programadores ($n = 2$), Analistas de TI ($n = 2$) e o restante não identificou a sua profissão ou função atual de trabalho.

Tabela 3 – Atribuição de Participantes aos Professores Avaliadores

Part. N°	Prof. A	Prof. B	Prof. C	Prof. D	Prof. E	Total de Avaliações
01	✓		✓	✓		03
02	✓		✓	✓		03
03	✓		✓	✓		03
04	✓		✓	✓		03
05	✓	✓			✗	02
06		✓			✓	02
07		✓			✓	02
08			✓	✓		02
09		✓			✓	02
10	✓				✓	02
11	✓			✓		02
12	✓	✓				02
13	✓				✓	02
14			✓	✓		02
15	✓			✓		02
16		✓			✗	01
17		✓			✗	01
18		✓			✗	01
19			✗	✓		01
20		✓	✗		✗	01
21		✓	✗			01
22			✗	✓		01

✓ Participante atribuído, parecer entregue pelo avaliador.

✗ Participante atribuído, parecer não-entregue.

Tabela 4 – Exemplo de par resposta / avaliação

<p>Part. N° 15 / Questão n° 1 - Você precisa adquirir um carro, quais os critérios que você utiliza para a compra?</p> <p><i>Faço economias e conto com estratégia para aquisição de recursos que virão em tempo determinado como restituição de imposto de renda e férias.</i></p>
<p>Prof. A / Avaliação para cada dimensão</p> <p>PA: <i>Atende parcialmente</i> DE: <i>Atende parcialmente</i> GE: <i>Atende totalmente</i> AB: <i>Atende totalmente</i> AV: <i>Atende totalmente</i></p> <p>Justificativa: <i>Apontou passos que são essenciais para esse participante, mas não é um conjunto de passos ordenados.</i></p>

As respostas foram então encaminhadas para os professores avaliadores. Cada professor-avaliador recebeu as respostas das provas de dez participantes, atribuídos aleatoriamente de modo a garantir que cada prova fosse avaliada por pelo menos dois professores, da maneira como ilustrado na Tabela 3. Foi dado então o prazo de 11 dias, de 03 e 14 de junho de 2020, para os professores avaliadores retornarem os seus pareceres. No entanto, devido à grande quantidade de esforço demandado pela tarefa, alguns professores (a saber, *Prof. C* e *Prof. E*) não retornaram todas as avaliações. Foi possível observar que todos os profissionais participantes tiveram as suas respostas apreciadas por pelo menos um dos avaliadores. Ao total, foram recebidas de volta 41 avaliações completas. A Tabela 4 apresenta um exemplo real de avaliação completa.

Considerando as avaliações de maneira isolada ($n=41$), é possível perceber que ficaram distribuídos de maneira equilibrada entre os quatro graus da escala (Tabela 5), apresentando apenas uma leve concentração nos graus “Não atende” e “Atende parcialmente”, demonstrando que não há um viés de avaliação significativo.

Numa primeira análise, foi possível observar um baixo nível de concordância entre os avaliadores, quando avaliadores os mesmos participantes. Considerando as respostas das questões individualmente, avaliadores distintos atribuíram exatamente a mesma avaliação (não é possível avaliar / não atende / atende parcialmente / atende totalmente) em apenas 46,7% ($n = 210$) de um universo de 450 casos possíveis (15 participantes foram avaliados por mais de um avaliador, nos cinco itens das seis questões). Este valor é uma evidência de que a subjetividade das habilidades de PC pode ser um desafio para a precisão do processo de avaliação. A Tabela 6 exibe uma avaliação mais específica por questão e dimensão do pensamento computacional. É possível observar uma leve tendência de maior concordância entre os avaliadores nas dimensões de habilidade de Pensamento algorítmico (PA) e Decomposição (DE), mas as dimensões restantes apresentam taxas de concordância evidentemente baixas.

Tabela 5 – Distribuição dos dados das avaliações (Dimensões / Questões)

Parecer	Dim.	Q1	Q2	Q3	Q4	Q5	Q6
Não é possível avaliar	PA	07% (3/41)	07% (3/41)	15% (6/41)	05% (2/41)	-	-
	DE	05% (2/41)	07% (3/41)	15% (6/41)	05% (2/41)	-	-
	GE	10% (4/41)	12% (5/41)	15% (6/41)	05% (2/41)	02% (1/41)	02% (1/41)
	AB	05% (2/41)	07% (3/41)	12% (5/41)	07% (3/41)	-	-
	AV	07% (3/41)	15% (6/41)	17% (7/41)	07% (3/41)	05% (2/41)	10% (4/41)
Não atende	PA	73% (30/41)	-	32% (13/41)	34% (14/41)	22% (9/41)	05% (2/41)
	DE	51% (21/41)	22% (9/41)	29% (12/41)	10% (4/41)	12% (5/41)	20% (8/41)
	GE	37% (15/41)	44% (18/41)	41% (17/41)	15% (6/41)	32% (13/41)	29% (12/41)
	AB	24% (10/41)	39% (16/41)	32% (13/41)	05% (2/41)	27% (11/41)	15% (6/41)
	AV	24% (10/41)	49% (20/41)	49% (20/41)	49% (20/41)	37% (15/41)	49% (20/41)
Atende parcialmente	PA	20% (8/41)	41% (17/41)	29% (12/41)	34% (14/41)	54% (22/41)	39% (16/41)
	DE	32% (13/41)	44% (18/41)	37% (15/41)	51% (21/41)	61% (25/41)	44% (18/41)
	GE	46% (19/41)	29% (12/41)	27% (11/41)	51% (21/41)	56% (23/41)	46% (19/41)
	AB	56% (23/41)	32% (13/41)	39% (16/41)	63% (26/41)	51% (21/41)	59% (24/41)
	AV	44% (18/41)	27% (11/41)	24% (10/41)	22% (9/41)	39% (16/41)	24% (10/41)
Atende totalmente	PA	-	51% (21/41)	24% (10/41)	27% (11/41)	24% (10/41)	56% (23/41)
	DE	12% (5/41)	27% (11/41)	20% (8/41)	34% (14/41)	27% (11/41)	37% (15/41)
	GE	07% (3/41)	15% (6/41)	17% (7/41)	29% (12/41)	10% (4/41)	22% (9/41)
	AB	15% (6/41)	22% (9/41)	17% (7/41)	24% (10/41)	22% (9/41)	27% (11/41)
	AV	24% (10/41)	10% (4/41)	10% (4/41)	22% (9/41)	20% (8/41)	17% (7/41)

Tabela 6 – Contagem de concordâncias entre duplas de avaliadores

Dimensão / Questão	PA	DE	GE	AB	AV	Valor médio por Questão:
Q1	73% (11/15)	67% (10/15)	33% (5/15)	33% (5/15)	47% (7/15)	51% (7,6/15)
Q2	53% (8/15)	40% (6/15)	40% (6/15)	40% (6/15)	47% (7/15)	44% (6,6/15)
Q3	60% (9/15)	40% (6/15)	40% (6/15)	47% (7/15)	33% (5/15)	44% (6,6/15)
Q4	53% (8/15)	47% (7/15)	47% (7/15)	47% (7/15)	20% (3/15)	43% (6,4/15)
Q5	60% (9/15)	67% (10/15)	47% (7/15)	73% (11/15)	40% (6/15)	57% (8,6/15)
Q6	40% (6/15)	60% (9/15)	33% (5/15)	47% (7/15)	27% (4/15)	41% (6,2/15)
Valor médio por Dimensão:	57% (8,5/15)	53% (8/15)	40% (6/15)	48% (7,1/15)	36% (5,3/15)	

Finalmente, foram coletas 128 textos das justificativas providas pelos avaliadores, as quais foram então submetidas ao processo de análise e síntese qualitativa. As justificativas apresentadas pelos avaliadores não necessariamente abordavam todas as cinco dimensões do pensamento computacional, e não necessariamente eram úteis para extrair os critérios de avaliação. Por exemplo, a justificativa “*descreveu passos em ordem*” foi associada à dimensão PA, justificativas do tipo “*não respondeu*” foram classificadas como não aplicáveis. Sendo assim, um total de 59% (76/128) sentenças contribuíram em pelo menos uma das cinco dimensões. A Tabela 7 ilustra o processo de codificação e categorização, e a Tabela 8 apresenta o conjunto de critérios resultantes. Como já era esperado, as justificativas abordaram mais frequentemente critérios relacionados à dimensão PA, enquanto dimensões como GE e DE obtiveram significativamente menos contribuições.

Para avaliar se um estudante desenvolveu ou não, por exemplo, a habilidade de AB, é preciso verificar se ele é capaz de apresentar uma solução a um dado problema de maneira a (AB1) apresentar informações relevantes sobre o contexto em que o problema está sendo resolvido, para que a partir daí possa (AB2) identificar os requisitos essenciais para a resolução do problema, e (AB3) apresentar uma solução simples, e (AB4) de maneira objetiva, ignorando detalhes irrelevantes para a avaliação. O mesmo raciocínio deste exemplo se aplica para as outras dimensões, de maneira que – um instrumento de avaliação elaborado com base neste conjunto de critérios é capaz inclusive avaliar precisamente quais são as especificidades dentro cada habilidade que foram efetivamente desenvolvidas pelos estudantes, e quais aspectos ainda precisam ser melhor trabalhados.

Tabela 7 – Exemplo da síntese qualitativa aplicado à dimensão AB

Passo 1: Identificação de aspectos relevantes semelhantes em diferentes justificativas	Passo 2: Categorização
<p>“Poderíamos dizer que há abstração por ter citado os <u>aspectos relevantes para a decisão</u>, embora isso seja induzido pela pergunta.” — Prof. E / Part. N° 7 / Q1 (AB: Atende parcialmente)</p> <p>“(…) <u>ausência de passos relevantes para a tarefa de plantar árvore</u> (...)” — Prof. D / Part. N° 2 / Q5 (AB: Atende parcialmente)</p> <p>“<u>identificou o principal</u>, que é separar por cor, mas não descreveu como lavar.” — Prof. A / Part. N° 10 / Q4 (AB: Atende totalmente)</p>	(AB) Identifica Requisitos Essenciais para Solução

Tabela 8 - Conjunto final de critérios para avaliação de habilidades de PC

PA	76% (58/76)	DE	9% (7/76)	GE	17% (13/76)	AB	33% (25/76)	AV	34% (26/76)
PA1. Lista uma sequência de instruções		DE1. Divide em subproblemas		GE1. Identifica Padrões		AB1. Contextualiza o problema		AV1. Identifica Critérios e Métricas	
PA2. Detalha das instruções		DE2. Compõe soluções para resolver o problema final		GE2. Discute a Generalidade da solução		AB2. Identifica Requisitos Essenciais para Solução		AV2. Pondera condições	
PA3. Apresenta Coerência Lógica						AB3. Apresenta Objetividade		AV3. Verifica Efetividade da solução	
PA4. Apresenta Completude						AB4. Apresenta Simplicidade			
PA5. Apresenta Corretude									

Finalmente, como exemplo para uso ou aplicação destes critérios, a rubrica que utilizamos para avaliação poderia apresentar as seguintes definições dos quatro níveis: **Não é possível avaliar** quando não é apresentada uma resposta à questão; **Não atende** quando nenhum dos critérios pode ser evidenciado na resposta apresentada; **Atende parcialmente** quando apenas um subconjunto dos critérios daquela dimensão pode ser evidenciado na resposta; e **Atende totalmente** quando todos os critérios daquela dimensão podem ser evidenciados. Outra forma de aplicação seria atribuindo notas e/ou pesos para a presença de cada um dos critérios, em cada uma das dimensões. Investigações futuras podem evoluir esses critérios ora apresentados no sentido de fundamentar sistemáticas completas de avaliação das competências específicas do PC.

5. Conclusão

Este trabalho teve como objetivo elicitar um conjunto de critérios para melhoria da precisão da avaliação das habilidades de pensamento computacional, a partir da prática. Para tanto, elaboramos uma réplica de prova contendo questões de pensamento computacional, coletamos repostas para estas provas com profissionais atuando na área de computação, e submetemos tais respostas à avaliação subjetiva de professores universitários experientes no ensino de PC.

Os dados expostos neste trabalho evidenciam o baixo nível de concordância entre diferentes avaliadores sobre todas as dimensões do pensamento computacional, confirmando que a precisão do processo de avaliação continua sendo um problema relevante para a prática. Com base nesses dados, realizamos uma síntese qualitativa que resultou em um conjunto explícito de critérios de avaliação para cada uma das dimensões do PC. Como resultado, chegamos a um conjunto de 16 critérios, que corroboram com outros trabalhos disponíveis na literatura, tal qual com os comportamentos do aprendiz descritos em Csizmadia [2014], apresentando, portanto,

uma validação empírica destes elementos. Por outro lado, os critérios das dimensões de DE, GE e AV possuem claras limitações devido à escassez de dados, fazendo com que critérios esperados (como “reutilização/importação de soluções pré-prontas de outros contextos para solução deste problema” na dimensão de GE) não aparecessem, levantando dúvidas sobre a saturação dos dados. Portanto, essas dimensões merecem ser melhor exploradas em pesquisas futuras.

Ao longo da execução deste trabalho, buscamos reduzir as ameaças à validade dos resultados, coletando dados concretos, e obtendo um conjunto de pareceres o mais próximo possível da prática do dia-a-dia dos professores. Compreendemos também que as tarefas, como elaboradas, ficaram trabalhosas, e isso acabou limitando a capacidade de respostas de alguns professores. Uma alternativa a este método seria o uso de dados de disciplinas executadas por estes professores, mas esta pesquisa foi realizada numa época de pandemia, em que disciplinas foram paralisadas e o acesso ao campo foi radicalmente limitado, inviabilizando este caminho. Estes resultados ainda não foram apresentados de volta e rediscutidos com os professores avaliadores, mas isso será feito numa interação futura deste trabalho, que faz parte do desenvolvimento da dissertação de mestrado do segundo autor. No entanto, entendemos que, mesmo diante de tais ameaças, os resultados alcançados por este trabalho têm o potencial de contribuir de maneira significativa em processos de avaliação em disciplinas de pensamento computacional.

No futuro, ainda é necessário investigar de fato qual é o impacto da adoção explícita destes critérios na elaboração de instrumentos de avaliação sobre a redução da variabilidade dos resultados destas avaliações. Esta verificação não fazia parte do escopo deste trabalho, portanto deve ser cuidadosamente considerada como uma limitação à validade de aplicação destes resultados alcançados aqui, e pode ser inclusive investigada por trabalhos futuros. Os métodos e ferramentas utilizados nesta pesquisa também representam uma contribuição no sentido de permitir que replicações futuras possam verificar ou melhorar os resultados alcançados neste trabalho. Todos os dados, formulários, instrumentos e análises apresentados neste trabalho encontram-se disponíveis com os autores para apreciação de terceiros, seja com fins à auditoria ou com fins à replicação dos estudos.

References

- AVILA, Christiano et al. Metodologias de Avaliação do Pensamento Computacional: uma revisão sistemática. In: Brazilian Symposium on Computers in Education (Simpósio Brasileiro de Informática na Educação-SBIE). 2017. p. 113.
- BLIKSTEIN, Paulo. O pensamento computacional e a reinvenção do computador na educação. Education & Courses, 2008.
- BRENNAN, Karen; RESNICK, Mitchel. New frameworks for studying and assessing the development of computational thinking. In: Proceedings of the 2012 annual meeting of the American educational research association, Vancouver, Canada. 2012. p. 25.
- CRUZES, Daniela S.; DYBA, Tore. Recommended steps for thematic synthesis in software engineering. In: 2011 international symposium on empirical software engineering and measurement. IEEE, 2011. p. 275-284.
- CSIZMADIA, Andrew et al. Computational thinking-A guide for teachers. 2015.

- DAVIES, Anna; FIDLER, Devin; GORBIS, Marina. Future work skills 2020. Institute for the Future for University of Phoenix Research Institute, v. 540, 2011.
- DE ARAUJO, Ana Liz Souto O.; ANDRADE, Wilkerson L.; GUERRERO, Dalton D. Serey. A systematic mapping study on assessing computational thinking abilities. In: 2016 IEEE frontiers in education conference (FIE). IEEE, 2016. p. 1-9.
- DE PAULA, B. H., VALENTE, J. A. e BURN, A. (2014). O uso de jogos digitais para o desenvolvimento do currículo para a Educação Computacional na Inglaterra. *Currículo sem Fronteiras*, v. 14, n. 3, p. 46-71.
- LEE, I.; MARTIN, F.; APONE, K. Integrating computational thinking across the K-8 curriculum. *ACM Inroads*, v. 5, n. 4, p. 64-71, 2014.
- RODRIGUEZ, Brandon et al. Assessing computational thinking in CS Unplugged Activities. In: *Proceedings of the 2017 ACM SIGCSE Technical Symposium on Computer Science Education*. 2017. p. 501-506.
- MORENO-LEÓN, Jesús; ROBLES, Gregorio; ROMÁN-GONZÁLEZ, Marcos. Comparing computational thinking development assessment scores with software complexity metrics. In: 2016 IEEE global engineering education conference (EDUCON). IEEE, 2016. p. 1040-1045.
- NATIONAL RESEARCH COUNCIL et al. Report of a workshop on the scope and nature of computational thinking. National Academies Press, 2010.
- RIBEIRO, Leila; FOSS, Luciana; CAVALHEIRO, Simone André da Costa. Entendendo o pensamento computacional. arXiv preprint arXiv:1707.00338, 2017.
- ROMÁN-GONZÁLEZ, Marcos. Computational thinking test: Design guidelines and content validation. In: *Proceedings of EDULEARN15 conference*. 2015. p. 2436-2444.
- ROMÁN-GONZÁLEZ, Marcos; MORENO-LEÓN, Jesús; ROBLES, Gregorio. Complementary tools for computational thinking assessment. In: *Proceedings of International Conference on Computational Thinking Education (CTE 2017)*, S. C Kong, J Sheldon, and K. Y Li (Eds.). The Education University of Hong Kong. 2017. p. 154-159.
- SEITER, Linda; FOREMAN, Brendan. Modeling the learning progressions of computational thinking of primary grade students. In: *Proceedings of the ninth annual international ACM conference on International computing education research*. 2013. p. 59-66.
- SELBY, Cynthia; WOOLLARD, John. Computational thinking: the developing definition. 2013.
- SELBY, Cynthia; DORLING, Mark; WOOLLARD, John. Evidence of assessing computational thinking. 2014.
- VALENTE, José Armando. Integração do pensamento computacional no currículo da educação básica: diferentes estratégias usadas e questões de formação de professores e avaliação do aluno. *Revista E-curriculum*, v. 14, n. 3, p. 864-897, 2016.
- WERNER, Linda et al. The fairy performance assessment: measuring computational thinking in middle school. In: *Proceedings of the 43rd ACM technical symposium on Computer Science Education*. 2012. p. 215-220.
- WING, Jeannette M. Computational thinking. *Communications of the ACM*, v. 49, n. 3, p. 33-35, 2006.
- ZHONG, Baichang et al. An exploration of three-dimensional integrated assessment for computational thinking. *Journal of Educational Computing Research*, v. 53, n. 4, p. 562-590, 2016.