

A Comparação da Realidade Mundial do Ensino de Programação para Iniciantes com a Realidade Nacional: Revisão sistemática da literatura em eventos brasileiros

Vinicius Ramos¹, Mateus Freitas², Maurício Galimberti²,
Antonio Carlos Mariani² e Raul Wazlawick²

¹Campus Araranguá — Universidade Federal de Santa Catarina (UFSC)

²Departamento de Informática e Estatística
Universidade Federal de Santa Catarina (UFSC)

{v.ramos, m.f.galimberti, antonio.c.mariani, raul.wazlawick}@ufsc.br

mateus.felipe@posgrad.ufsc.br

Abstract. *This paper presents a systematic literature review (SLR) about algorithms and programming teaching for beginner students in Brazilian higher education. Those courses are called CS1 (Computer Science 1) for short. The review was conducted on papers published in the following Brazilian conferences: SBIE, WIE, WEI and WAAAP between the years 2001 and 2014. This work searches for empirical evidences about the artefacts that influence success/failure rates in CS1 courses. Furthermore, the results of this review are compared to international results. From 394 papers selected 49 were analysed, but only 7 presented failure rates before and after experimentation. The failure rate on those papers was observed to be reduced from 45.6 to 32.6% after experimentation.*

Resumo. *Este artigo apresenta uma revisão sistemática da literatura (RSL) sobre o ensino de programação e algoritmos para alunos iniciantes do ensino superior brasileiro, chamados CS1 (Computer Science 1). A RSL foi realizada à partir da revisão dos artigos publicados nos eventos SBIE, WIE, WEI e WAAAP entre os anos de 2001 a 2014. Este trabalho busca evidências empíricas dos artefatos que influenciam as taxas de sucesso/reprovação nos cursos CS1. Além disso, procurou-se comparar resultados desta revisão com os resultados apresentados internacionalmente. Do total de 394 artigos pré-selecionados, restaram 49, dos quais apenas 7 artigos apresentam taxas de reprovação de alunos antes e depois da realização do(s) experimento(s), reduzindo essa taxa, em média, de 45,6% para 32,6%.*

1. Introdução

O ensino de programação é parte integrante dos currículos de referência da Sociedade Brasileira de Computação (SBC). As disciplinas de introdução à programação são incluídas, geralmente, no primeiro semestre dos cursos. Existem diferentes nomes para este componente curricular. Neste artigo optamos por chamá-lo de CS1 (*Computer Science 1*) para usarmos a nomenclatura mundialmente conhecida.

[Bennedsen and Caspersen 2007] realizaram pesquisa sobre as taxas de sucesso/reprovação nos cursos CS1 em busca de dados empíricos que validassem os rumores do alto índice de reprovações nesses cursos. A taxa de reprovação é de 33%, onde cursos maiores têm pesos maiores, e de 28% para valores absolutos.

[Watson and Li 2014] buscaram novos indícios das taxas de reprovação dos alunos dos cursos CS1 apontadas por [Bennedsen and Caspersen 2007]. Os autores encontraram como média das taxas de reprovação de 32,3%, muito próximos ao anterior. Os autores avaliaram 52 artigos em sua revisão sistemática da literatura (RSL).

[Vihavainen et al. 2014] tentam compreender, de forma quantitativa à partir de uma RSL sobre práticas pedagógicas no ensino de cursos CS1 e motivados pela taxa média de reprovação de 33% apresentada por [Bennedsen and Caspersen 2007], quais as estratégias de ensino influenciam a taxa de reprovação dos alunos nos cursos CS1. A média das taxas de reprovação estava em 38,6% e passou para 25,6% após as intervenções.

Porém, a realidade brasileira do ensino de CS1 pode ser diferente de outros países, onde grandes turmas com até 200 ou 300 alunos são formadas para palestras e depois os alunos são alocados a tutores para aulas práticas em laboratório. No Brasil, o Ministério da Educação não recomenda turmas com mais de 50 alunos e as aulas de laboratório normalmente são conduzidas pelo próprio professor da disciplina. Assim, diferenças podem existir entre a efetividade das práticas testadas no exterior e as do nosso país.

A partir deste contexto, buscamos evidências empíricas dos artefatos (ferramentas, estratégias, etc) que influenciam as taxas de sucesso ou reprovações nas instituições de ensino superior brasileiras à partir da revisão de artigos publicados no Simpósio Brasileiro de Informática na Educação (SBIE)¹, Workshop de Informática na Escola (WIE)² e Workshop sobre Educação em Computação (WEI)³ entre os anos de 2001 a 2014 e o II Workshop de Ambientes de apoio à Aprendizagem de Algoritmos e Programação (WAAAP)⁴ no ano de 2008. Portanto, o nosso objetivo com esse estudo é comparar os resultados obtidos nesta RSL (Revisão Sistemática da Literatura) com os resultados apresentados internacionalmente.

Dispomos as seções deste artigo da seguinte maneira. A Seção 3 apresenta o método utilizado para a realização da RSL deste trabalho. Na Seção 4, mostramos uma análise detalhada sobre a qualidade dos artigos encontrados, algumas ameaças à validade da RSL e trabalhos relacionados a este. A Seção 5 descreve os resultados obtidos com esta RSL. Por fim, apresentamos algumas conclusões à partir dos resultados encontrados e comentamos sobre possíveis implicações e trabalhos futuros.

¹O SBIE é o evento mais importante da área de informática na educação no Brasil, por isso a sua escolha.

²O WIE é um workshop consolidado e que existe desde 1995, a partir de 2011 ele passou a fazer parte do SBIE.

³O WEI tem por objetivo debater diferentes temas relacionados ao ensino de computação e informática. É um evento anual produzido pela Sociedade Brasileira de Computação (SBC) e é um dos principais fóruns brasileiros que visa discutir a melhoria do processo de ensino e de aprendizagem em Computação.

⁴O WAAAP teve apenas duas edições e a não inclusão da primeira edição deve-se ao fato de não encontrarmos referências digitais para os seus trabalhos.

2. Trabalhos Relacionados

Um relato sobre o ensino de programação no Brasil é feito por [Pereira Júnior and Rapkiewicz 2004]. Os autores apresentam uma visão geral dos trabalhos publicados no SBIE e WEI entre os anos de 1999 e 2004, quantificam e categorizam cada um dos trabalhos em: ferramenta, estratégia ou ferramenta com estratégia. Este é um dos primeiros relatos panorâmicos sobre o ensino de programação no Brasil. Neste trabalho, os autores apresentam resultados incipientes sobre o ensino de programação, mostrando a carência deste tipo de trabalho na área da Computação. Um outro ponto importante é o predomínio de ferramentas nos trabalhos avaliados.

[da Silva Ribeiro et al. 2012] apresenta uma visão geral sobre o ensino de programação e algoritmos baseados nos eventos nacionais SBIE, WIE e WEI entre 2004 e 2011. O principal objetivo dos autores é mostrar as principais linhas utilizadas no desenvolvimento de software à partir do tema “ensino de algoritmos e programação”. Além das categorias propostas por [Pereira Júnior and Rapkiewicz 2004], os autores sintetizam em subcategorias os artigos encontrados e referenciam um ou dois desses trabalhos para cada uma delas.

[Aureliano and Tedesco 2012a] apresentam uma RSL do ensino de programação para iniciantes focada no SBIE e WIE entre os anos de 2002 e 2011. As autoras apresentam um panorama geral sobre as produções científicas no Brasil, buscando enumerar as instituições de ensino envolvidas, os níveis de escolaridade e os tipos de artefatos propostos nas publicações. Como principais resultados, as autoras apontam para uma predominância de instituições do Sul e Sudeste, sendo que as pesquisas foram feitas em sua maioria no ensino superior e aproximadamente 65% dos artigos apresentam ferramentas de software. Em suas conclusões, as autoras apontam para a falta de análise da qualidade das publicações, um dos pontos que difere-se da nossa pesquisa.

3. Método Utilizado

Este trabalho está estruturado de acordo com a proposta de [Kitchenham 2004] para uma revisão sistemática da literatura. Neste estudo, o objetivo é encontrar evidências empíricas sobre os artefatos produzidos que influenciam positiva ou negativamente as taxas de reprovação de cursos CS1 em instituições de ensino superior (IES) brasileiras. Desta forma, é possível fazer comparações com os artigos recentes que apresentam estas evidências no cenário internacional [Vihavainen et al. 2014, Watson and Li 2014].

Neste contexto, a nossa questão de pesquisa está definida como “Existem artefatos (softwares, metodologias, materiais de apoio, etc) que influenciam nas reprovações dos cursos CS1 nas IES brasileiras?”. A partir desta questão, apresentamos outras questões, que chamaremos de questões norteadoras, que visam nos auxiliar na extração e sintetização dos dados dos artigos encontrados nesta RSL, são elas:

- Quais artefatos produzidos influenciam nas aprovações/reprovações dos cursos CS1 nas IES brasileiras?
- A configuração (número de alunos, horas de aula, quantidade de tutores/monitores, horas em laboratório, avaliações) dos cursos CS1 nas IES brasileiras influenciam nas aprovações/reprovações desses cursos?

3.1. Mecanismos e string de busca

Neste artigo, fazemos uma análise dos artigos publicados no SBIE, WIE, WEI e WAAAP, de 2001 a 2014 em busca de responder à questão de pesquisa apresentada no início desta seção. Utilizamos os mecanismos de busca desses eventos, exceto o WAAAP, pois ele não possui um desses. Desta forma, todos os 15 artigos aceitos no II WAAAP foram incluídos no estudo primário.

Para a realização das buscas, criamos uma *string* formada por palavras-chave e por operadores lógicos E (AND) e OU (OR), aqui apresentada: (*Ensino OR Disciplina OR Curso OR Causa OR Fator OR Desempenho OR Influência OR Motivo*) AND (*Informática OR Computação OR Programação OR Algoritmo OR CSI*).

3.2. Critérios de inclusão e exclusão

Adotamos uma abordagem em quatro etapas para a busca. Na primeira etapa, realizamos a busca usando a nossa *string* nos mecanismos de buscas do SBIE e do WIE e incluímos os 15 artigos do II WAAAP no conjunto de artigos primários. Foram identificados 127 artigos do WIE, 175 artigos do SBIE, 77 artigos do WEI e os 15 artigos do WAAAP, num total de 394 artigos.

Na segunda etapa, lemos os títulos, resumos e palavras-chave de cada um dos artigos. Após a leitura, sobrou um total de 137 artigos. Para esta leitura, utilizamos os critérios de inclusão (CI) e os critérios de exclusão (CE) apresentados na Tabela 1

Tabela 1. Critérios de inclusão e exclusão de artigos primários

Critérios de Inclusão	Critérios de Exclusão
CI1. Publicações relacionadas ao ensino introdutório de programação ou algoritmos para alunos do ensino superior brasileiro.	CE1. Excluir publicações nas quais o foco não é o ensino introdutório de programação ou algoritmos.
CI2. Publicações entre os anos de 2001 e 2014.	CE2. Artigos que apresentam propostas, ideias, ferramentas ou estratégias sem apresentar resultados, avaliações ou validações.
CI3. Ao encontrar artigos com os mesmos autores e relatos parecidos, mas em eventos diferentes, apenas um (o mais recente) foi incluído.	CE3. Exclusão de artigos resumidos, poster, <i>survey</i> , revisão sistemática e mapeamento sistemático.
	CE4. Artigos duplicados ou parecidos.

Na terceira etapa, foram lidas as seções de resultados e/ou de conclusão. A maioria dos artigos (73 no total) excluídos encaixam-se no critério de exclusão 2 apresentado na Tabela 1. Restando apenas 5 artigos do WIE, 11 do SBIE, 27 do WEI e 6 do WAAAP. Num total de 49 artigos referenciados em <https://goo.gl/dw5Fku>.

Na quarta etapa, todos os 49 artigos foram lidos integralmente. Na próxima seção, damos prosseguimento à RSL de forma a extrair dados de qualidade dos artigos.

4. Análise da Revisão Sistemática da Literatura

Além dos critérios de inclusão e exclusão que devem estar claros em uma RSL, [Kitchenham 2004] aponta a importância de se verificar a qualidade dos artigos selecionados e reduzir o viés da pesquisa. Nesta seção, apresentamos os critérios de qualidade usados neste estudo e a presença ou não destes critérios nos artigos avaliados.

4.1. Qualidade dos Artigos Selecionados

[Kitchenham 2004] considera importante avaliar a qualidade dos artigos primários, pois ela traz maior clareza aos critérios de inclusão e exclusão dos artigos, pode explicar se diferenças na qualidade de estudos traz diferenças no resultado deles, é uma maneira de pesar a importância de trabalhos individuais quando estudos são consolidados, pode guiar nas interpretações dos resultados e na influência das inferências e podem guiar em futuras publicações. Um problema, ainda segundo a autora, é que não há consenso quando o assunto é qualidade. Apesar disso, a autora traz referências apontando que a qualidade deve minimizar o viés e aumentar a validade interna e externa da pesquisa.

Além dos dados principais dos 49 artigos avaliados neste trabalho, como, o nome do revisor, data da extração, referência bibliográfica e algumas possíveis notas, esses artigos foram catalogados em uma planilha com o objetivo de apresentar claramente as suas principais características, que seguem: objetivos do estudo, tipo do estudo, número de participantes envolvidos, como os experimentos foram realizados, linguagem de programação usada, nome da(s) ferramenta(s)/software, estratégias de ensino usadas, resultados, conclusões e dificuldades encontradas, motivos ou fatores que deram início ao estudo e avaliação da qualidade do estudo. A planilha completa pode ser encontrada em: <http://goo.gl/MqkLLf>.

A Tabela 2 mostra os critérios de qualidade a serem avaliados nos artigos primários desta RSL [Kitchenham 2004, Dybå and Dingsøyr 2008, da Silva et al. 2014].

Dividimos os critérios de qualidade levando em consideração os 6 tópicos que devem estar presentes em pesquisas empíricas de acordo com as diretrizes de [Kitchenham et al. 2002]: 1. contexto do experimento; 2. planejamento do experimento; 3. condução do experimento e coleta dos dados; 4. análise; 5. apresentação dos resultados; e 6. interpretação dos resultados.

A Tabela 3 mostra um consolidado dos critérios de qualidade pelos artigos selecionados (a tabela completa pode ser encontrada em <http://goo.gl/TLofBD>).

Se por um lado a Tabela 3 mostra altos percentuais referentes aos critérios de qualidade CQ1, CQ2 e CQ3 (87,8%, 77,7% e 79,6%, respectivamente), o que nos mostra que as pesquisas são baseadas em dados empíricos, com base em relatórios ou na opinião de especialistas, possuem objetivos bem definidos e procuram definir claramente o contexto onde a pesquisa foi realizada. Por outro lado, nos mostra claramente uma baixa preocupação com a descrição, planejamento, análise e apresentação dos resultados dos experimentos realizados nos artigos.

De fato, os outros critérios de qualidade não chegam nem a 50% dos artigos analisados. O CQ8, que versa sobre o rigor na análise dos dados coletados, obteve o pior valor dentre todos. Apenas 5 (10,2%) dos trabalhos avaliados apresentam rigor na sua análise dos resultados. Utilizamos como base para o nosso entendimento de “análise com

Tabela 2. Critérios de Avaliação de Qualidade

Critérios	Descrição
CQ1	O estudo está baseado em pesquisas empíricas ou em relatos de experiência com base em relatórios ou na opinião de especialistas?
CQ2	Existe uma definição clara dos objetivos da pesquisa?
CQ3	O contexto da pesquisa está adequadamente descrito?
CQ4	O planejamento da pesquisa atende aos objetivos?
CQ5	A estratégia de extração de dados foi adequada aos objetivos da pesquisa?
CQ6	Haviam critérios bem definidos (grupo de controle, elementos comparativos, turmas novas) com o qual pudesse comparar tratamentos?
CQ7	Os dados foram coletados de forma a abordar as questões de pesquisa?
CQ8	A análise dos dados foi suficientemente rigorosa?
CQ9	A relação pesquisador-participantes foi considerada em grau adequado?
CQ10	Existe uma indicação clara dos resultados?

Tabela 3. Consolidado dos critérios de qualidade encontrados nos artigos selecionados e seus respectivos percentuais

CQ1	CQ2	CQ3	CQ4	CQ5	CQ6	CQ7	CQ8	CQ9	CQ10
43	38	39	17	17	11	15	5	12	22
87,8%	77,6%	79,6%	34,7%	34,7%	22,4%	30,6%	10,2%	24,5%	44,9%

rigor” as 5 diretrizes apresentadas por [Kitchenham et al. 2002]: 1. especificação de todo e qualquer procedimento usado para controlar múltiplos testes; 2. considerar o uso de “análise às cegas”; 3. realizar análises minuciosas (identificação e tratamento de valores atípicos ou de observações que influenciam a análise); 4. assegurar que os dados não violam as restrições aplicadas pelos testes a serem utilizados; e 5. aplicar procedimentos de controle de qualidade apropriados para a verificação dos resultados.

Outros dois critérios, CQ6 e CQ9, apresentam percentuais muito baixos: 22,4% e 24,5%, respectivamente. Ambos descrevem a forma como os experimentos foram conduzidos e os dados coletados. Estes valores não invalidam os experimentos, tão pouco assumimos que esse controle de qualidade não tenha sido aplicado durante a realização deles, apenas verificamos que estes critérios são pouco abordados nos artigos, dificultando a sua repetição. Os critérios CQ5 e CQ7 possuem valores baixo (34,7% e 30,6%, respectivamente) e muito próximos, pois ambos estão intimamente relacionados. Ambos os critérios versam sobre o planejamento, a condução e a coleta dos dados do experimento.

A falta desses indicativos de qualidade estão diretamente relacionados à replicação do experimento por outros pesquisadores. Sabemos que este problema é uma realidade na computação, mais especificamente na Engenharia de Software como afirma [Kitchenham et al. 2002]. Entretanto, para mudarmos este cenário, é importante cons-

cientizarmos os pesquisadores da área do importante papel do planejamento, execução e apresentação dos resultados dos experimentos conforme os rigores científicos.

5. Resultados

Apresentamos nesta seção os principais resultados encontrados à partir da RSL de cursos CS1 nas IES brasileiras. Buscamos evidências empíricas dos artefatos que influenciam as taxas de sucesso/reprovação nas IES brasileiras, para que seja possível compararmos estes artefatos com os resultados apresentados internacionalmente, como em [Vihavainen et al. 2014, Watson and Li 2014]. Tentaremos responder, também, à nossa questão de pesquisa e às questões norteadoras apresentadas na Seção 3.

Procuramos por indícios das taxas de reprovação antes e/ou depois das intervenções apresentadas nos artigos. Encontramos apenas 7 artigos com essas características [Barbosa et al. 2014, Campos 2010, Deters et al. 2008, Gondim et al. 2009, Píccolo et al. 2010, Piva Jr and Freitas 2011, Valaski and Paraiso 2012]. Nesses artigos, a média das taxas de reprovações antes das intervenções é de aproximadamente 45,6%. [Campos 2010] apresenta os dados mais baixos de reprovação: 37% para alunos novatos e 23% para alunos repetentes. [Deters et al. 2008] apresentam a maior taxa de reprovação antes dos experimentos: 56,41%.

Os 7 artigos não apresentam dados brutos de reprovação após a realização das intervenções. Alguns dividem os dados entre alunos novatos e alunos repetentes, dificultando o nosso cálculo da média. Entretanto, considerando apenas os dados de alunos novatos, quando existem (quando não há divisão na apresentação dos dados, incluímos o total), a média da taxa de reprovação cai para 32,6%, aproximando-se dos estudos realizados por [Vihavainen et al. 2014, Watson and Li 2014].

Para [Barbosa et al. 2014] e [Píccolo et al. 2010] as taxas de reprovação após a aplicação de suas estratégias de ensino são mantidas em torno de 50%. Apesar de os índices de reprovação não diferirem antes e após as intervenções, [Barbosa et al. 2014] mostra uma diferença significativa no uso da linguagem Python em detrimento à linguagem C para o ensino em cursos CS1 (excluindo-se os repetentes), tendo, inclusive, uma pequena redução nas taxas de reprovação de 48% para 40%.

Dentre os outros artigos, [Deters et al. 2008] afirmam que com a aplicação de sua metodologia de ensino, 80% dos alunos obtiveram sucesso na sua segunda experiência com a disciplina.

[Piva Jr and Freitas 2011] apontam uma redução das reprovações de 45,6% para 32,7% após a aplicação de suas estratégias de ensino em dois anos seguidos, num total de 4 turmas de CS1.

Para [Gondim et al. 2009] a taxa de reprovação caiu de 50% para 10% após a utilização de um ambiente visual de programação 3D conhecido como ALICE. Já [Valaski and Paraiso 2012] mostraram-se surpresos com esse resultado e descrevem em seu artigo: “Para nossa surpresa, obtivemos resultados diferentes dos apresentados em [Gondim et al. 2009]. A utilização do ambiente interativo 3D Alice não levou a substancial melhora na aprovação e aprendizado na disciplina de programação.”

[Silva et al. 2011] apresentam uma metodologia que aumentou em 80% o índice de aprovação nas turmas que usaram esta metodologia.

O uso de jogos como fator de motivação é destaque em [de Jesus et al. 2014]. A taxa de reprovação, após a inclusão de uma metodologia que utiliza jogos para ensinar, é de 20% para os repetentes e de 50% para os novatos. É importante ressaltar que [da Silva et al. 2014] afirmam que em sua RSL “97% dos estudos relatam que a utilização de jogos é uma ferramenta eficaz de ensino e aprendizagem de programação.”.

[Digiampietri et al. 2010] também usam jogos como fator motivacional no ensino de programação. Em seus estudos, os autores apontam uma redução na taxa de evasão de 20% para menos de 10%. Para os autores a utilização dos jogos aumentou o interesse dos alunos no curso superior que frequentam e, portanto, os autores atribuem à utilização de jogos no ensino de programação a essa redução da evasão.

Segundo a perspectiva de [Pereira Júnior and Rapkiewicz 2004], dividimos os artigos avaliados em categorias e quantidades: a. ferramentas (9); b. estratégias (18); c. ferramentas com estratégias (14); d. materiais didáticos⁵; e e. outros⁶

Não foi possível comparar diretamente estes resultados com os resultados apontados por [Pereira Júnior and Rapkiewicz 2004] e [Aureliano and Tedesco 2012a], pois os nossos critérios de inclusão e exclusão são distintos dos que eles propunham. Entretanto, acredita-se que a diferença entre essas revisões e a nossa deve-se ao fato de que muitas ferramentas apresentadas nos eventos que revisamos não apresentam dados de avaliação da ferramenta em si, pois um dos critérios de exclusão (CE2. Tabela 1) da nossa RSL é a não inclusão destes dados.

Dentre os artigos que apresentam dados sobre as turmas que foram submetidas às intervenções, todos possuem número de alunos reduzidos, menores que 50. Esta é uma das diferenças entre as turmas apresentadas nos artigos encontrados por [Vihavainen et al. 2014] e a realidade brasileira. Outra diferença entre a realidade brasileira e a internacional é que a maioria das aulas são realizadas em laboratório.

6. Conclusões e Trabalhos Futuros

Realizamos uma revisão sistemática da literatura (RSL) com o objetivo de buscar evidências empíricas dos artefatos que influenciam as taxas de sucesso/reprovação nas instituições de ensino superior brasileiras. Comparamos os resultados encontrados nesta RSL com dados apresentados por [Vihavainen et al. 2014, Watson and Li 2014] e os artigos referenciados por eles, para verificar as diferenças entre o panorama de ensino-aprendizagem de programação para iniciantes no Brasil e fora dele.

Após as etapas de busca, a aplicação dos critérios de inclusão e exclusão (veja na Seção 3.2) e análise, restaram 49 artigos que foram incluídos nesta revisão. Desses artigos, apenas 7 fazem referência direta às taxas de aprovação e reprovação antes e depois das intervenções apresentadas pelos autores. Isto demonstra que, em um espaço de 14 anos, ainda temos muito o que discutir nessa área, principalmente quando falamos de artefatos que podem reduzir os problemas encontrados no processo de ensino-aprendizagem de programação e algoritmos para iniciantes.

⁵Materiais didáticos: objetos de aprendizagem, vídeos, apostilas, livros, entre outros

⁶Outros: revisão de exercícios, revisão da taxonomia de Bloom para criação de atividades de avaliação, revisão de ferramentas para o ensino e correlações entre a disciplina de programação, outras disciplinas iniciais e o processo seletivo

Os critérios de qualidade (CQ) que apresentamos na Seção 4.1 indicam diversos pontos onde os autores da área de ensino de programação, engenharia de software, desenvolvimento de materiais, entre tantos outros que participaram dos artigos lidos nesta revisão têm grandes dificuldades. Baseando-se nas diretrizes de [Kitchenham et al. 2002], consideramos que apenas 7 (14,29%) dos artigos preenchem 7 ou mais critérios de qualidade (a tabela completa pode ser encontrada em <http://goo.gl/TLoFBD>). Ainda com relação aos CQ, [Aureliano and Tedesco 2012b] é o único artigo que possui todos os CQ apresentados em nossa pesquisa. Com o intuito de aumentar a qualidade dos artigos da área, recomendamos a elaboração dos experimentos baseados nas diretrizes de [Kitchenham et al. 2002] e na elaboração e identificação dos possíveis estudos empíricos apresentados por [Easterbrook et al. 2008].

Como trabalhos futuros, pretendemos identificar os pesquisadores no Brasil que têm como principal área de atuação o ensino de algoritmos e programação com publicações na área em busca de criar uma comunidade para compartilhamento de materiais, experiências, softwares e resultados que ajudem a reduzir os problemas da área.

Pretendemos, ainda, criar o portal de ensino de algoritmos e programação, com o objetivo de criar uma comunidade de ensino, onde poderemos compartilhar artefatos e incentivar a melhoria da qualidade no processo de ensino-aprendizagem da área.

Referências

- [Aureliano and Tedesco 2012a] Aureliano, V. C. O. and Tedesco, P. C. d. A. R. (2012a). Ensino-aprendizagem de programação para iniciantes: uma revisão sistemática da literatura focada no SBIE e WIE. In *Anais do Simpósio Brasileiro de Informática na Educação*, volume 23.
- [Aureliano and Tedesco 2012b] Aureliano, V. C. O. and Tedesco, P. d. A. (2012b). Avaliando o uso do scratch como abordagem alternativa para o processo de ensino-aprendizagem de programação. In *Congresso da Sociedade Brasileira de Computação*, volume 32, pages 1–10.
- [Barbosa et al. 2014] Barbosa, A. d. A., Ferreira, D. Í., and Costa, E. B. (2014). Influência da linguagem no ensino introdutório de programação. In *Anais do Simpósio Brasileiro de Informática na Educação*, volume 25, pages 612–621.
- [Bennedsen and Caspersen 2007] Bennedsen, J. and Caspersen, M. E. (2007). Failure Rates in Introductory Programming. *SIGCSE Bull.*, 39(2):32–36.
- [Campos 2010] Campos, R. (2010). Metodologia erm2c: Para melhoria do processo de ensino-aprendizagem de lógica de programação. In *XVIII Workshop sobre Educação em Computação*.
- [da Silva et al. 2014] da Silva, T. R., Medeiros, T. J., and Aranha, E. H. d. S. (2014). Jogos digitais para ensino e aprendizagem de programação: uma revisão sistemática da literatura. In *Anais do Simpósio Brasileiro de Informática na Educação*, volume 25, pages 692–701.
- [da Silva Ribeiro et al. 2012] da Silva Ribeiro, R., Brandão, L. d. O., and Brandão, A. A. (2012). Uma visão do cenário nacional do ensino de algoritmos e programação: uma proposta baseada no paradigma de programação visual. In *Anais do Simpósio Brasileiro de Informática na Educação*, volume 23.
- [de Jesus et al. 2014] de Jesus, Â. M., Gonçalves, D. A. S., and Ferreira, L. A. C. (2014). Aplicação de desenvolvimento de jogos digitais como um meio de motivação em di-

- ferentes níveis de ensino de computação. In *Anais do Workshop de Informática na Escola*, volume 20.
- [Deters et al. 2008] Deters, J. I., Silva, J. d., Miranda, E., and Fernandes, A. (2008). O desafio de trabalhar com alunos repetentes na disciplina de algoritmos e programação. In *Workshop de Ambientes de apoio à Aprendizagem de Algoritmos e Programação*.
- [Digiampietri et al. 2010] Digiampietri, L. A., Kropiwiec, D. D., and Silva, R. A. (2010). O uso de jogos como fator motivacional em cursos de computação.
- [Dybå and Dingsøyr 2008] Dybå, T. and Dingsøyr, T. (2008). Empirical studies of agile software development: A systematic review. *Information and software technology*, 50(9):833–859.
- [Easterbrook et al. 2008] Easterbrook, S., Singer, J., Storey, M.-A., and Damian, D. (2008). Selecting empirical methods for software engineering research. In *Guide to advanced empirical software engineering*, pages 285–311. Springer.
- [Gondim et al. 2009] Gondim, H. W. A., Ambrósio, A. P., and Costa, F. M. (2009). Uma experiência no ensino de algoritmos utilizando ambientes visuais de programação 3d. In *Workshop sobre Educação em Computação (WEI), Bento Gonçalves-RS*.
- [Kitchenham 2004] Kitchenham, B. (2004). Procedures for performing systematic reviews. *Keele, UK, Keele University*, 33(2004):1–26.
- [Kitchenham et al. 2002] Kitchenham, B. A., Pfleeger, S. L., Pickard, L. M., Jones, P. W., Hoaglin, D. C., El Emam, K., and Rosenberg, J. (2002). Preliminary guidelines for empirical research in software engineering. *Software Engineering, IEEE Transactions on*, 28(8):721–734.
- [Pereira Júnior and Rapkiewicz 2004] Pereira Júnior, J. C. R. and Rapkiewicz, C. E. (2004). O processo de ensino-aprendizagem de fundamentos de programação: Uma visão crítica da pesquisa no Brasil. In *WEI-Workshop sobre Educação em Computação*, pages 19–21.
- [Píccolo et al. 2010] Píccolo, H. L., Sena, V. d. F., Nogueira, K. B., da Silva, M. O., and Maia, Y. A. (2010). Ambiente interativo e adaptável para ensino de programação. In *Anais do Simpósio Brasileiro de Informática na Educação*, volume 1.
- [Piva Jr and Freitas 2011] Piva Jr, D. and Freitas, R. L. (2011). Estratégias para melhorar os processos de abstração na disciplina de algoritmos. In *XXXI Congresso da Sociedade Brasileira de Computação, Natal/RN. Anais do XIX Workshop sobre Educação em Computação. Natal, RN: SBC*, pages 1488–1497.
- [Silva et al. 2011] Silva, G. C., Ré, R., Kawamoto, A., and Schwerz, A. (2011). Uma experiência na aplicação de práticas de apoio no ensino-aprendizado de algoritmos. In *Anais do Workshop de Informática na Escola*, volume 1, pages 1378–1381.
- [Valaski and Paraiso 2012] Valaski, J. and Paraiso, E. C. (2012). Limitações da utilização do Alice no ensino de programação para alunos de graduação. In *Anais do Simpósio Brasileiro de Informática na Educação*, volume 23.
- [Vihavainen et al. 2014] Vihavainen, A., Airaksinen, J., and Watson, C. (2014). A Systematic Review of Approaches for Teaching Introductory Programming and Their Influence on Success. In *Proceedings of the Tenth Annual Conference on International Computing Education Research, ICER '14*, pages 19–26, New York, NY, USA. ACM.
- [Watson and Li 2014] Watson, C. and Li, F. W. (2014). Failure Rates in Introductory Programming Revisited. In *Proceedings of the 2014 Conference on Innovation & Technology in Computer Science Education, ITiCSE '14*, pages 39–44, New York, NY, USA. ACM.