

BOAS PRÁTICAS




Dúvidas estatísticas

Software que refaz cálculos de artigos e detecta erros movimentados a comunidade de pesquisadores da área da psicologia

Em agosto de 2016, um conjunto de cerca de 50 mil artigos científicos do campo da psicologia passou pelo escrutínio de um software capaz de detectar inconsistências estatísticas, desenvolvido por pesquisadores da Universidade de Tilburg, na Holanda. Batizado de *statcheck*, o programa refaz cálculos e observa se os resultados são robustos e correspondem ao que está publicado no manuscrito. Em metade dos *papers* foi detectado algum tipo de problema, desde erros de digitação e simples arredondamentos até resultados equivocados com potencial para comprometer conclusões de estudos. Os achados dessa gigantesca verificação foram comunicados de forma automática por e-mail aos autores de cada artigo e divulgados no PubPeer, plataforma on-line na qual qualquer usuário pode escrever comentários sobre *papers* já publicados e apontar eventuais falhas, num tipo de avaliação por pares realizado depois da divulgação do artigo.

O caráter da auditoria é inédito, tanto por ter sido feita por computador quanto pelo volume de dados conferido – praticamente todos os pesquisadores em psicologia que publicaram *papers* nos últimos 20 anos passaram pelo crivo do *statcheck*. A divulgação dos resultados provocou ondas de choque. A Sociedade de Psicologia da Alemanha publicou uma declaração, no dia 20 de outubro, reclamando da forma como a comunicação dos resultados foi feita. Segundo o texto, muitos pesquisadores ficaram incomodados com a exposição sem que tivessem tido chance de se defender. “Vários colegas estão profundamente preocupados com o fato de que, obviamente, é muito difícil remover um comentário no PubPeer após um erro ‘detectado’ pelo *statcheck* revelar-se um falso positivo”, consta no texto divulgado.

Num tom mais elevado, Susan Fiske, professora da Universidade de Princeton e ex-presidente da norte-americana Associação de Ciências



Psicológicas, classificou como “uma nova forma de assédio” o trabalho de “polícia” que investiga dados de pesquisa de forma voluntarista. “Me senti um pouco assustado e exposto”, disse ao jornal britânico *The Guardian* o psicólogo alemão Mathias Kauff, que recebeu um e-mail do *statcheck* avisando que havia inconsistências num artigo que escreveu em 2013 sobre multiculturalismo e preconceito, publicado no *Personality and Social Psychology Bulletin*. Ele afirma que os erros eram fruto de arredondamentos que não comprometiam as conclusões.

Muitos artigos na área de psicologia utilizam testes estatísticos padronizados, cujos resultados precisam ser averiguados. O *statcheck* identifica e inspeciona os testes que calculam os valores de p , uma medida que representa a probabilidade de o efeito observado dever-se ao acaso e não aos fatores que estão sendo estudados. Um valor de p menor ou igual a 0,05 é utilizado frequentemente como um indicador de significância estatística, pois sugere que os resultados são robustos.

Há, de fato, evidências de que o software ainda não está maduro e alardeia problemas que não chegam a ser erros estatísticos. Em artigo depositado no repositório ArXiv, Thomas Schmidt, professor de psicologia experimental da Universidade de Kaiserslautern, na Alemanha, criticou a qualidade da análise que o *statcheck* fez em dois artigos de sua autoria. Segundo ele, o software encontrou 35 resultados estatísticos potencialmente incorretos, mas apenas cinco continham incongruências que, de acordo com o autor, não comprometiam os resultados finais.

A metodologia utilizada pelo software já era conhecida desde 2015, quando um artigo sobre o assunto foi publicado no site da revista *Behavior Research Methods*, assinado pela



Chris Hartgerink, o estudante de doutorado que submeteu *papers* de psicologia ao *statcheck*

estudante de doutorado Michèle Nuijten e colegas do Centro de Metapesquisa da Escola de Ciências Sociais e do Comportamento da Universidade de Tilburg. No *paper*, o grupo mostrou que metade dos 16.695 artigos analisados pelo software apresentou algum tipo de inconsistência em suas análises estatísticas e 12% deles traziam conclusões comprometidas por erros. “O *statcheck* pode ser uma ferramenta de apoio à revisão por pares. A revista *Psychological Science*, por exemplo, já adotou o software para procurar inconsistências estatísticas nos manuscritos que recebe”, disse Michèle à *Pesquisa FAPESP*.

A iniciativa de analisar os 50 mil artigos e tornar públicos os resultados no PubPeer foi do estudante de doutorado Chris Hartgerink, de 25 anos. Segundo ele, a intenção foi gerar benefícios imediatos para o campo da psicologia, que não seriam alcançados se apenas resultados gerais fossem divulgados. O fato de também detectar falsos positivos e erros sem importância, disse o pesquisador, não compromete esse objetivo. Hartgerink e o professor Marcel van Assen tentam agora desenvolver outro tipo de software, capaz de detectar se um artigo científico contém dados

fabricados. Para testar a eficiência do método, a dupla pediu a colegas para enviar versões de seus *papers* com dados deliberadamente alterados, que estão sendo avaliados.

Entre os pesquisadores de psicologia, há também quem considere o *statcheck* uma ferramenta útil para melhorar a qualidade das publicações científicas. Simine Vazire, pesquisadora do Departamento de Psicologia da Universidade da Califórnia em Davis, prevê que os autores de artigos nessa área serão ainda mais cuidadosos com suas análises estatísticas agora que sabem da existência de um programa capaz de identificar descuidos, erros e fraudes.

A Universidade de Tilburg, onde o programa foi desenvolvido, foi o cenário de um escândalo de má conduta científica. Em setembro de 2011, a instituição demitiu um de seus mais produtivos pesquisadores, o professor de psicologia social Diederik Stapel, acusado de fraudar mais de 30 artigos científicos ao longo de oito anos – uma investigação provou que ele fabricava dados, enganava coautores e até mesmo intimidava quem desconfiasse dele (ver *Pesquisa FAPESP* nº 190).

Chris Hartgerink foi aluno de graduação de Stapel e tinha o professor como uma espécie de mentor – fora inclusive seu assistente de pesquisa. Ficou, na época, desorientado. “Era uma figura inspiradora e foi o responsável pelo meu entusiasmo em fazer pesquisa”, disse ao jornal *The Guardian*. A experiência amarga do escândalo levou parte do grupo de pesquisadores que investigou suas fraudes a montar o Centro de Metapesquisa, interessado no estudo de má conduta científica. Hartgerink uniu-se ao grupo em 2013, em seu projeto de doutorado sobre métodos para detectar a fabricação de dados de pesquisa. ■

Antídotos contra notícias falsas

Um dos temas que movimentaram a reunião anual da Associação Americana para o Avanço da Ciência, realizada em Boston em fevereiro, foi a proliferação de notícias falsas na internet. Uma palestra de Kevin Elliott, professor de ética da Universidade do Estado de Michigan, mostrou que o fenômeno, que ganhou destaque no mundo político dos Estados Unidos desde a última campanha eleitoral, é um antigo conhecido da ciência. Elliott mencionou os estudos enviesados patrocinados durante décadas pela indústria do tabaco para mascarar os efeitos deletérios do cigarro para a saúde, além de casos recentes, como a manipulação de testes de emissão de poluentes promovida pela Volkswagen, descoberta em 2015.

Um bom antídoto contra as notícias científicas falsas, segundo o pesquisador, é manter algum ceticismo quando o responsável pela pesquisa tem algum interesse direto no resultado favorável. O mais seguro, quando se deseja saber o que a ciência tem a dizer sobre um problema específico, é consultar os relatórios temáticos que sociedades científicas respeitadas costumam produzir. “Assim, evitam-se interpretações excêntricas”, afirmou Elliott.

Já a apresentação de Dominique Brossard, professora da Universidade de Wisconsin-Madison, propôs que, na divulgação de fatos científicos, é tênue a linha que separa a fraude pura e simples e os efeitos do jornalismo de má qualidade. A falsa informação, disse, é disseminada com o intuito de enganar e influenciar pessoas. Ela citou um estudo que realizou com um aluno da Universidade Yale sobre as notícias bizarras publicadas num tabloide sensacionalista distribuído em supermercados nos Estados Unidos, como recém-nascidos com 15 quilos de peso, ataques de alienígenas e insetos gigantes.

Segundo o estudo, a maioria das informações é inventada, mas uma parte é composta por histórias reais inusitadas, sem que o leitor possa separar uma coisa da outra.

Já o jornalismo científico de má qualidade produz situações nebulosas. Dominique citou um estudo que ganhou repercussão nas redes sociais, segundo o qual a cafeína preveniria o câncer, mas que se baseava num ensaio com apenas 10 indivíduos. “Os jornalistas não são treinados para aferir a validade de um estudo. Eles tentam enfatizar o lado humano da notícia em manchetes como: ‘Novos estudos trazem esperança para familiares de vítimas do mal de Alzheimer’”, comentou, de acordo com o serviço de notícias EurekAlert. Como isso gera expectativas, espalha-se nas redes sociais.

Dominique sugeriu três estratégias para enfrentar o problema. A primeira é uma convocação para que cientistas se disponham a explicar melhor o que estão fazendo e a ajudar jornalistas a avaliar os achados científicos. A segunda é envolver



ILUSTRAÇÃO JÚLIA CHEREM RODRIGUES

instituições científicas no monitoramento nas redes sociais de notícias falsas envolvendo suas pesquisas, divulgando esclarecimentos sempre que necessário. A terceira é convencer as ferramentas de busca na internet a retirar de seus registros referências a trabalhos científicos que sofreram retratação.

Yuans em troca de artigos

O sociólogo Jeroen Huisman, professor da Universidade de Ghent, da Bélgica, recebeu um e-mail de um representante da Universidade Zhengzhou, da China, perguntando se ele teria interesse em passar uma temporada na instituição como professor visitante. Huisman ficou curioso e pediu mais detalhes. Para sua surpresa, recebeu a minuta de um contrato que previa o pagamento de 300 mil yuans, o equivalente a R\$ 135 mil, caso ele produzisse três *papers* em revistas indexadas e declarasse a universidade chinesa como sua instituição de origem. O contrato previa apenas duas visitas a

Zhengzhou, capital da província de Henan. Ele rejeitou a oferta. “Não parecia ilegal, mas era evidentemente antiético. Tratava-se apenas de uma transação financeira”, disse. A revista *Times Higher Education* procurou a universidade chinesa, que não quis se pronunciar. Rui Yang, professor da Faculdade de Educação da Universidade de Hong Kong, disse à revista que esse tipo de contrato é comum em universidades chinesas – ele próprio já rejeitou ofertas. “Algumas universidades precisam de bons artigos em inglês para não ficarem mal nos processos de avaliação realizados no país.”