توهم کفایت؛ آیا دانستن نام سوگیری‌های شناختی می‌تواند درمان نادانی در دوران ما باشد؟

رامین سعادت

چکیده

این مقاله، دومین از سه‌گانهٔ «توهم کفایت»، این باور رایج را به آزمون می‌گذارد که دانستن نام یک سوگیری شناختی، نیمی از راه اصلاح آن است. با تکیه بر چارچوب چهار شرطی ویلسون و برکه (۱۹۹۴) و شواهد رفتاری از آزمایش‌های لنگرسازی، نشان داده می‌شود که آگاهی از وجود یک سوگیری، حتی وقتی صریحاً به فرد هشدار داده شود، به‌تنهایی خطای قضاوت را اصلاح نمی‌کند؛ شکافی که در این مقاله «فرانادانی» (Meta-ignorance) نامیده می‌شود. در ادامه، مقاله نشان می‌دهد این فرانادانی در سطح فردی باقی نمی‌ماند: برچسب سوگیری در منازعات هویتی به ابزار بی‌اعتبارسازی رقیب بدل می‌شود، و در اقتصاد توجه رسانه‌ها از زمینهٔ تجربی و روش‌شناختی دقیقش جدا می‌گردد؛ همچنانکه سرنوشت اثر دانینگ-کروگر در گفتمان عمومی نشان می‌دهد. با این حال، مقاله نشان می‌دهد ادعایش اصلاح‌کننده است، نه نقض‌کننده: شواهد فراتحلیلی موجود نشان می‌دهند آموزشی که تمرین، بازخورد و کاربرد واقعی را در بر بگیرد، می‌تواند اثری کوچک اما معنادار بر کاهش سوگیری داشته باشد، هرچند انتقال آن به تصمیم‌گیری واقعی هنوز قطعی نیست و طراحی محیط تصمیم، به‌عنوان مسیری مستقل، می‌تواند مکمل یا جایگزین آن باشد. نتیجهٔ کلی این است که دانستن نام، آگاهی می‌سازد اما اصلاح نمی‌کند؛ و اصلاح واقعی، تنها از رهگذر تمرین و بازخورد به دست می‌آید.

واژگان کلیدی: سوگیری‌های شناختی؛ نقطهٔ کور سوگیری؛ فرانادانی؛ آموزش اصلاح سوگیری؛ اثر دانینگ-کروگر؛ برچسب‌زنی هویتی؛ طراحی محیط تصمیم؛ تفکر انتقادی

مقدمه و جایگاه این مقاله

این مقاله، دومین، از سه‌گانه‌ای با عنوان مشترک «توهم کفایت» است. مقاله نخست نشان داد شناخت «مغالطات منطقی»، هرچند مفید، به‌تنهایی برای مقابله با نادانی دوران ما کافی نیست. این مقاله همان پرسش را دربارهٔ حوزهٔ محتوایی دیگری می‌پرسد: «سوگیری‌های شناختی». مغالطات از سنت منطق می‌آیند و سوگیری‌ها از سنت روان‌شناسی شناختی؛ که هر دو، به لحاظ روش‌شناسی در زیر چتر روش‌شناختی «تفکر انتقادی» قرار می‌گیرند. به بیان دیگر، این دو حوزه‌هایی محتوایی‌اند که «مهارت‌های تفکر انتقادی» روی آنها اعمال می‌شوند، نه اینکه خودشان یکی از آن مهارت‌ها به شمار آیند.

مقاله سوم این مجموعه، به شکافی خواهد پرداخت که در خود ادبیات تفکر انتقادی میان مهارت و گرایش به کاربرد آن شناخته شده است.

پیش از ادامه، اما لازم است مرز دیگری نیز روشن شود؛ مرز میان «سوگیری‌های شناختی» (Cognitive Biases)، که موضوع این مقاله است، و آنچه گاه در زبان عمومی، و حتی در برخی منابع، زیر عنوان عام‌تر «خطاهای شناختی» (Cognitive Errors) گرد هم می‌آید.

واقعیت این است که عنوان عام «خطای شناختی»، در عمل، غالباً فهرست‌هایی را هم در بر می‌گیرد که از سنتی کاملاً جداگانه می‌آیند؛ سنت «تحریفات شناختی» (Cognitive Distortions)، که در حوزهٔ «روان‌درمانی شناختی-رفتاری» قرار دارد. موضوعی که نخستین‌بار با بک (Beck, 1976) صورت‌بندی مفهومی شد. پس از او، برنز آن را به فهرستی دقیقاً شمارش‌شده (ده مورد) بدل کرد (Burns, 1980) و بعدها در بازبینی ۱۹۹۹ اصلاحش کرد (Burns, 1999).

با گذشت زمان اما، منابع و کاربست‌های عامه‌پسند بعدی، اغلب فهرستی گسترده‌تر و ناهم‌خوان با آن ارائه می‌دهند؛ مواردی مانند: ذهن‌خوانی، فاجعه‌سازی، بایدها، شخصی‌سازی، استدلال احساسی، و مواردی از این دست

(Sage, Keppens & Yannakoudakis, 2025). فهرستی که با آنکه نامش به «سوگیری شناختی» شبیه است، اما از جنس دیگری است: ابزاری بالینی برای شناسایی الگوهای فکریِ مرتبط با افسردگی و اضطراب در مراجع درمانی، نه یافته‌ای پژوهشی دربارهٔ «خطای قضاوت در افراد عادی». این مقاله به‌عمد وارد آن حوزه نمی‌شود؛ نه از سر کم اهمیت شمردن آن حوزه، بلکه چون آمیختن آن با سوگیری‌های شناختی زیر یک عنوان مشترک، همان خلطی را تکرار می‌کند که بخش ۳ این مقاله (در موردی دیگر)، نقد می‌کند؛ یعنی جداشدن یک برچسب علمی از زمینهٔ دقیق و سنت پژوهشی‌اش. موضوع این مقاله، به‌طور مشخص، سوگیری‌های شناختی به معنای دقیق آن، یعنی: «سنت پژوهش قضاوت و تصمیم‌گیری» است؛ همان سنتی که بعداً خواهیم دید چگونه توسط کانمن، تورسکی، و ادامه‌دهندگانشان شکل داده شده است.

اکنون که محدودهٔ کار و مرزبندی‌های مفهومی در این حوزه روشن شد، می‌توانیم به موضوع این مقاله و رسالت آن در دوران کنونی بپردازیم.

نوشتن مقاله‌ای درباب جنبه‌های مختلف موضوع سوگیری‌های شناختی از آنجا موجه می‌نماید که به نظر می‌رسد در سال‌های اخیر، دانستن نام ده‌ها نمونه از آنچه در بین مردم و علاقه‌مندان به حوزه تفکر انتقادی «سوگیری‌های شناختی» خوانده می‌شود، به یکی از پرتقاضاترین ژانرهای محتوای عمومی بدل شده است. نمونهٔ روشن این تقاضا، «چکیدهٔ سوگیری‌های شناختی» باستر بنسون (Benson, 2016) است که «جان مانوجیان سوم» اندکی بعد آن را به نگاشتی تصویری با عنوان «کدکس سوگیری‌های شناختی» بدل کرد؛ نگاشتی که بیش از ۱۸۰ سوگیری را در چهار دسته سامان می‌دهد و از آن زمان به‌طور گسترده بازنشر و در محیط‌های آموزشی و کسب‌وکاری استفاده شده است (Manoogian, 2016).

در واقع این باور عامیانه که «دانستن نام یک الگوی خطا نیمی از راه غلبه بر آن است»، امروزه آن‌قدر رایج است که نامی هم برایش ساخته شده و لوری سانتوس و تامار گندلر (هر دو از دانشگاه ییل) آن را «مغالطهٔ جی‌آی‌جو» (G.I. Joe Fallacy) نامیده‌اند (Santos & Gendler, 2014)؛ در اشاره به شعار کارتون جی‌آی‌جو در دههٔ ۱۹۸۰ که هر قسمت را با «دانستن نیمی از نبرد است» تمام می‌کرد. بعداً در سال ۲۰۲۱ کریستال و سانتوس، همین پدیده را با عنوان «پدیدهٔ جی‌آی‌جو» (G.I. Joe phenomena) به‌طور مستقیم بررسی کرده‌اند (Kristal & Santos, 2021). توضیح کریستال و سانتوس از این پدیده بر این اساس است که: آگاهی فراشناختی از یک سوگیری، به‌تنهایی، محدودیت‌های چندانی در کاهش عملی آن ندارد.

این مقاله دقیقاً همین باور و تعریف را، برای سوگیری‌های شناختی، به آزمون می‌گذارد و تز دقیق آن این است:

دانش مفهومی دربارهٔ یک سوگیری (نام، تعریف، مثال‌های آن) با ظرفیت تشخیص و اصلاح آن در لحظهٔ عمل یکی نیست؛ و شواهد موجود نشان می‌دهند که انتقال از اولی به دومی، بدون تمرین، بازخورد، و گاه تغییر محیط تصمیم، «محدود و نامطمئن» است. این ادعا دو چیز را رد نمی‌کند: نه اینکه چنین دانشی «بی‌فایده» است، و نه اینکه «هیچ آموزشی اثر ندارد». آنچه رد می‌شود، این باور است که صرف دانستن نام، به‌خودی‌خود، کفایت می‌کند.

همچنین در این مقاله، به‌جای آنکه فهرستی از سوگیری‌ها را ارائه دهیم، چند نمونهٔ شاخص (لنگر، نقطهٔ کور سوگیری، و اثر دانینگ-کروگر) را به‌عنوان مطالعات موردی برای آزمودن یک ادعای عمومی‌تر دربارهٔ کل این مقوله بررسی خواهیم کرد.

اما در همینجا لازم است پیشاپیش به یک پرسش پاسخ داده شود: سوگیری شناختی، به‌عنوان یک «خطای قضاوت»، چه نسبتی با نادانی به معنای عام آن (موضوع این مجموعه) دارد؟

از نگاه این مقاله، پاسخ را باید در «شواهد رفتاری» جست، نه در گزارش خودِ افراد از «نقطهٔ کور سوگیری» (Bias Blind Spot) (این باور رایج که خود را کمتر از دیگران سوگیرانه می‌بینیم)؛ زیرا می‌دانیم حتی وقتی به فرد صریحاً گفته می‌شود در معرض یک سوگیری مشخص قرار دارد نیز، این «آگاهی» به‌تنهایی قضاوتش را اصلاح نمی‌کند (بخش ۱)، و وقتی آموزشی فراتر از «دانستن نام»، طراحی می‌شود، باز هم میزان انتقال آن به تصمیم‌گیری واقعی نامطمئن و غیر قابل اعتماد است (بخش۴).

بر اساس همین دو یافتهٔ مستند رفتاری، که در ادامه به تفصیل خواهد آمد، ادعای مقاله این است که چه بسا فرد از وجود یک «محدودیت در قضاوت خودش» کاملاً آگاه باشد ولی همچنان نتواند آن را در لحظهٔ عمل اصلاح کند. همین شکاف میان «آگاهی مفهومی» و «اصلاح رفتاری» است که در این مقاله «فرانادانی» (Meta-ignorance) نامیده می‌شود؛ نه به معنی «نبود دانش

دربارهٔ جهان بیرون»، بلکه «نارسایی دانش در ترجمه‌شدن به عمل». تزی که به هیچیک از تبیین‌های رقیب برای مکانیسم نقطهٔ کور (بخش ۱) به تنهایی متکی نیست.

این «فرانادانی»، یعنی ندانستنِ خودِ ناتوانی در اصلاح سوگیری، اما تنها در سطح فردی و شناختی باقی نمی‌ماند. همین برچسب که قرار بود ابزار خودشناسی باشد، در میدان نزاع‌های هویتی به سلاحی علیه رقیب بدل می‌شود، و در اقتصاد توجه رسانه‌ها از بار علمی‌اش تهی و به اتهامی سطحی تقلیل می‌یابد؛ پیش از آنکه به شواهدی برسیم که نشان می‌دهند آموزش واقعی، برخلاف صرف دانستن نام، در کجا و تا چه حد اثر می‌کند.

برای بررسی این موضوعات، در ادامه بخش‌های زیر را خواهیم داشت:

بخش ۱: چارچوب نظری، یعنی چرا دانستن نام یک سوگیری برای اصلاح آن کافی نیست؛

بخش ۲: تبدیل برچسب سوگیری به سلاحی در نزاع هویتی؛

بخش ۳: جدا افتادن برچسب از معنای علمی‌اش در اقتصاد توجه و رسانه؛

بخش ۴: شواهدی که نشان می‌دهند آموزش واقعی، برخلاف صرف دانستن نام، چه زمانی و تا چه حد مؤثر است.

---

۱. چارچوب نظری: چرا دانستن نام کافی نیست

فرانادانی، یا شکافی که میان دانستن و اصلاح برشمردیم، نیازمند صورت‌بندی دقیق‌تری است؛ باید دقیقاً روشن کنیم چه شرط‌هایی باید برآورده شوند تا یک «قضاوت سوگیرانه» اصلاح شود، و دانستن نام یک سوگیری، دقیقاً در کدام‌یک از آنها نقشی ندارد.

در همین زمینه، ویلسون و برکه (Wilson & Brekke, 1994) برای «اصلاح هر قضاوت سوگیرانه» چهار شرط را لازم می‌دانند:

آگاهی از وجود سوگیری؛

انگیزه برای اصلاح آن؛

دانش دربارهٔ جهت و اندازهٔ تأثیرش؛

و توان کنترل یا اعمال یک راهبرد اصلاحی.

همچنانکه می‌بینیم، از میان موارد این فهرست، دانستن نام یک سوگیری در بهترین حالت، فقط شرط اول را تأمین می‌کند. سه شرط دیگر (انگیزه، دانش دربارهٔ اندازه و جهت خطا در آن لحظهٔ خاص، و توان اعمال راهبرد) نه از دانستن نام نتیجه می‌شوند و نه با آن تضمین می‌شوند.

اما برای آنکه این ادعا قابل آزمون باشد، باید از همین‌جا دو اصطلاح را به دقت تعریف کرد:

نخست اینکه، مراد از «دانستن نام» در این مقاله، مواجههٔ صرف با نام، تعریف، و مثال‌های یک سوگیری است؛ چیزی که با خواندن یک متن، دیدن یک اینفوگرافیک، یا شنیدن یک سخنرانی به دست می‌آید، بدون تمرین فعال، بازخورد اصلاحی، یا تجربهٔ کاربرد آن در یک تصمیم واقعی.

دوم، مراد از «آموزش» هر مواجههٔ دیگری است که حداقل یکی از این سه عنصر: «تمرین، بازخورد، یا کاربرد در یک موقعیت شبیه‌سازی‌شده یا واقعی» را داشته باشد.

این مرزبندی، و دقیق کردن مفاهیم، از این جهت ارزشمند است که اگر پژوهشی نشان دهد مواجههٔ صرف با نام و تعریف یک سوگیری، بدون هیچ‌یک از آن سه عنصر (تمرین، بازخورد، کاربرد)، خطای قضاوت واقعی را کاهش می‌دهد، تز این مقاله نقض شده است.

شاهد مؤید این نکته، در مطالعهٔ سال ۱۹۹۶ ویلسون و همکاران آمده است (Wilson, Houston, Etling & Brekke, 1996)؛ در این پژوهش، به شرکت‌کنندگان عددی «لنگر» (برگرفته از شمارهٔ شناساییِ خودشان) داده شد و از آنها خواسته شد آن را با تعداد پزشکان فهرست‌شده در دفترچهٔ تلفن شهرشان مقایسه و سپس برآورد خود را ارائه دهند. پیش از این مقایسه، به گروهی از شرکت‌کنندگان صریحاً هشدار داده شد که عددی در ذهنشان می‌تواند بر پاسخشان به پرسش بعدی اثر بگذارد؛ یعنی شرط اول مدل ویلسون و برکه (آگاهی از وجود سوگیری) را به‌طور کامل و مستقیم تأمین کردند. «نتیجه این بود که برآوردهای همهٔ گروه‌های

لنگرشده (چه هشدارداده‌شده و چه نشده) هم‌چنان به‌طور معنادار ($p< ۰/۰۵$) با گروه کنترلِ بدون‌لنگر تفاوت داشت؛ و در عین حال، میان خودِ گروه‌های لنگرشده، هشدار داده شده و هشدار داده نشده، هیچ تفاوت معناداری دیده نشد. یعنی هشدار، به‌تنهایی، نه اثر لنگر را از میان برد و نه حتی اندکی تضعیف کرد.

از نگاه روانشناسی شناختی نیز می‌توان این شکاف را در «چارچوب دوسیستمی کانمن» (Kahneman, 2011) دقیق‌تر دید. به این معنی که «دانستن نام یک سوگیری» محصول سیستم ۲ است (کند، صریح، نیازمند تلاش)، اما خود «وقوع سوگیری» در سیستم ۱ رخ می‌دهد (سریع، خودکار)، پیش از آنکه سیستم ۲ فرصت مداخله پیدا کند. از این دیدگاه، حتی وقتی سیستم ۲ نام سوگیری را می‌داند، «تنبلی ذاتی‌اش» باعث می‌شود در لحظهٔ عمل، غالباً به سیستم ۱ میدان بدهد.

نمونهٔ موازی این تعبیر از پدیدهٔ شکاف را روزنبلیت و کیل با عنوان «توهم عمق تبیین» (Illusion of Explanatory Depth) نشان دادند (Rozenblit & Keil, 2002). در آزمایش اصلی، افراد فکر می‌کردند مکانیسم وسایل روزمرهٔ ساده (از سیفون توالت تا قفل سیلندری) را می‌فهمند، تا وقتی خواسته شد آن را گام به گام توضیح دهند؛ آنگاه عمق فهمشان به‌شدت کمتر از آنچه گمان می‌کردند آشکار شد.

در نمونه‌ای دیگر از همین پژوهش‌ها درمورد توهم در قلمرو مهارت‌های پیچیده، لاوسون (Lawson, 2006) آن را برای کارکرد دوچرخه، و فرنباخ و همکاران (Fernbach, Rogers, Fox & Sloman, 2013) برای مواضع سیاستی آزمایش کردند و به نتایج مشابه رسیدند. به همین قیاس می‌توان گمان برد که دانستن نام یک سوگیری، «توهمی از همان جنس» می‌سازد؛ یعنی فرد می‌پندارد چون نامش را می‌داند، بر آن مسلط است، بدون آنکه هرگز این تسلط را در یک موقعیت واقعی آزموده باشد.

در اینجا البته باید روشن کرد که این گام، قیاسی است، نه نتیجه‌ای که مستقیماً از مطالعات بالا (که موضوعشان مهارت‌های دیگر بود، نه خودِ سوگیری‌های شناختی) به دست آمده باشد؛ اما سازوکار مشترک هر دو یکسان است: توهم تسلط، برآمده از مواجههٔ صرف، بدون آزمون عملی.

اما افزون بر اینها، باید توجه داشت که حتی همان شرط سادهٔ اول یعنی آگاهی از وجود سوگیری نیز همیشه به‌سادگی تأمین نمی‌شود، تا چه رسد به تأثیرگذاری آن بر اصلاح قضاوت‌های ما. موضوعی که برای نمونه، در پژوهش پرونین، لین و راس نشان داده شد (Pronin, Lin & Ross, 2002)؛ وقتی این پژوهشگران از شرکت‌کنندگان خواستند میزان سوگیری خود و میزان سوگیری یک فرد متوسط آمریکایی را در چندین سوگیری شناخته شده ارزیابی کنند، افراد به‌طور پیوسته، باور خود را کمتر از دیگران سوگیرانه ارزیابی کردند. نویسندگان این الگو را «نقطهٔ کور سوگیری» (Bias Blind Spot) نامیدند که پیشتر بیان کردیم. تبیین اصل پژوهشگران درباره این مفهوم نقطهٔ کور سوگیری، در آن مقاله این بود که: ادراک خود فرد از واقعیت بی‌واسطه و عینی است، و بنابراین هرکس با آن ادراک مخالف باشد، باید مرجح یا مغرض باشد.

البته چند سال بعد، خود پرونین و همکارانش تبیین دیگری به نام «توهم درون‌نگری» (Introspection Illusion) را نیز صورت‌بندی کردند (Pronin, Gilovich & Ross, 2004؛ Pronin & Kugler, 2007). بر اساس این تبیین، درون‌نگری (مرور افکار و نیات خود) گواهی قابل‌اعتماد برای سنجش سوگیری خود است، در حالی‌که دربارهٔ دیگران فقط به رفتار بیرونی‌شان دسترسی داریم؛ و چون سوگیری غالباً ناخودآگاه عمل می‌کند، درون‌نگری آن را در خودمان آشکار نمی‌کند، هرچند رفتار آن را در دیگران لو می‌دهد.

اما امروز دیگر نمی‌توان یکی از این دو تبیین را نهایی دانست؛ زیرا کارهای دیگری نیز در این زمینه انجام گرفته است. برای نمونه، ماندل، کالینز، واکر، فوگلسانگ و ریسکو، برای بررسی این موضوع، و علاوه بر تبیین «توهم درون‌نگری»، دو تبیین دیگر را نیز مستقیماً آزمودند (Mandel, Collins, Walker, Fugelsang & Risko, 2022).

یکی تبیین «پختگی شناختی» (cognitive sophistication)، که می‌گوید نقطهٔ کور اصلاً خطا نیست، بلکه بازتاب واقعیت است: افراد با توانایی شناختی و گرایش‌های تفکر بهتر، واقعاً کمتر از دیگران سوگیرانه قضاوت می‌کنند، پس خودارزیابیِ پایین‌ترشان درست است، نه اشتباه.

دیگری تبیین «فرآیندهای گفت‌وگویی» (conversational processes) که نقطهٔ کور را محصول نحوهٔ طرح پرسش می‌داند، نه خطایی واقعی در خودشناسی. به این معنی که وقتی به شرکت‌کننده گفته می‌شود فلان سوگیری در میان مردم عمومیت دارد و

سپس از او خواسته می‌شود «فرد متوسط» را ارزیابی کند، او بر اساس قواعد عادی گفت‌وگو نتیجه می‌گیرد این تعمیم به «دیگران» مربوط است، نه لزوماً به خودش.

نتایج ماندل و همکارانش در این مطالعه مشاهده کردند که:

تبیین «فرآیندهای گفت‌وگویی» بیشترین پشتیبانی را گرفت، چون اکثر شرکت‌کنندگان دیگران را بالاتر از نقطهٔ میانهٔ مقیاس ارزیابی کردند، در حالی‌که خودارزیابی‌شان نوسان بیشتری داشت و به‌ندرت به همان نقطه می‌رسید.

تبیین «پختگی شناختی» تأیید نشد، چون در آزمون‌های عینی سوگیری، میزان سوگیری واقعی افراد با سطح پختگی شناختی‌شان رابطه‌ای نداشت.

همچنین تبیین «توهم درون‌نگری» نیز تأیید نشد، چون الگویی که این تبیین پیش‌بینی می‌کرد (تکیهٔ بیشتر بر درون‌نگری هنگام قضاوت دربارهٔ خود، در برابر تکیهٔ بیشتر بر گرایش‌های رفتاری هنگام قضاوت دربارهٔ دیگران) در داده‌ها دیده نشد.

افزون بر این، نمرهٔ توهم درون‌نگری همبستگی ضعیفی با میزان کلی نقطهٔ کور داشت، و پس از کنترل توانایی شناختی و گرایش به «تفکر فعالانه و باز به شواهد مخالف» (Actively Open-minded Thinking/AOT)، دیگر پیش‌بینی‌کنندهٔ معناداری نبود.

در اینجا یک نکتهٔ ظریف وجود دارد که باید روشن شود.

اگر نامتقارنی خود-دیگر تا حدی زاییدهٔ «نحوهٔ طرح پرسش» باشد (همان‌طور که تبیین فرآیندهای گفت‌وگویی می‌گوید)، ممکن است این‌طور به نظر برسد که اصل پدیدهٔ نقطهٔ کور هم زیر سؤال می‌رود، و در نتیجه، تکیهٔ این مقاله بر آن نیز.

اما این‌طور نیست، زیرا مناقشه بر سر مکانیسم نقطهٔ کور (اینکه چرا این نامتقارنی رخ می‌دهد) مستقل است از شواهد رفتاری‌ای که تز این مقاله بر آنها تکیه دارد. وست، مزرو و استانوویچ در پژوهش خودشان نشان دادند که توانایی شناختی بالاتر نقطهٔ کور را کم نمی‌کند، و در برخی تحلیل‌ها، حتی اندکی افزایش نیز می‌دهد (West, Meserve & Stanovich, 2012). اسکوپلیتی و همکاران نیز نشان دادند افراد با نقطهٔ کور بالاتر توصیهٔ دیگران را کمتر می‌پذیرند و در وظایف تصمیم‌گیری عینی ضعیف‌تر عمل می‌کنند (Scopelliti et al., 2015).

این دو یافته، مستقل از هر یک از آن سه تبیین رقیب به دست آمده‌اند و نشان می‌دهند فرد، صرف‌نظر از آنکه نقطهٔ کورش از کدام مکانیسم ناشی شود، همچنان نمی‌تواند آن را با دانستن نامش اصلاح کند. به بیان دیگر، مناقشه بر سر مکانیسم، تصویر دقیق‌تری از نقطهٔ کور می‌دهد، بدون آنکه نتیجهٔ رفتاری‌ای که تز این مقاله به آن متکی است را تغییر دهد.

اما جدا از مناقشه بر سر مکانیسم، باید دربارهٔ اصل پدیدهٔ نقطهٔ کور سوگیری نیز سخن گفت.

واقعیت این است که همین پدیدهٔ «نامتقارنی خود-دیگر» (که یک نمونهٔ خاص آن نقطهٔ کور سوگیری است) در پژوهش‌های بعدی نیز به‌طور نسبتاً پایدار بازتولید شده است. برای نمونه، یک بازتولید پیش‌ثبت شده در برزیل (Seda et al., 2024)، اثر را (در جمعیتی خارج از نمونه‌های معمول آمریکای شمالی و دانشجویی) تأیید کرد، اما همین بازتولید یک محدودیت در نظریهٔ اصلی را هم نشان داد: برخلاف پیش‌بینی پرونین و همکاران، «نامتقارنی خود-دیگر» فقط برای «سوگیری» دیده نشد؛ برای «ضعف‌های شخصی» عمومی‌تر هم دیده شد.

همچنین بازتولید دیگری، با نمونه‌هایی از هنگ‌کنگ و آمریکا، همین نامتقارنی را در طرحی متفاوت تأیید کرده است (Chandrashekar et al., 2021)، که این بار آن را به «باورهای فرد دربارهٔ اختیار» هم پیوند داد. نتیجهٔ محتاطانه‌ای که می‌توان از این پژوهش‌ها گرفت این است که: اصل پدیده از تبیین‌های نظری‌اش مطمئن‌تر است، و مناقشهٔ جاری امروز، بر سر چرایی آن است، نه بر سر اصل وجودش.

نکتهٔ مهم در اینجا این است که باید به یاد داشت این توصیف از نقطهٔ کور، به‌معنای «کوری دائمی و درمان‌ناپذیر» نیست. در بخش ۴ خواهیم دید که در شرایط آموزشی مشخص (یعنی طبق تعریف بالا، آموزشی که تمرین و بازخورد دارد، نه صرف دانستن نام) خود نقطهٔ کور هم می‌تواند کاهش پیدا کند؛ آنچه این بخش نشان می‌دهد، محدود به این ادعا است که صرف دانستن نام یک سوگیری، بدون چنان آموزشی، نقطهٔ کور را کم نمی‌کند.

البته یک محدودیت نیز در خود این چارچوب وجود دارد که باید گفته شود، و آن هشدار گرد گیگرنزر (Gigerenzer, 2018) است به آنچه او آن را «سوگیری‌انگاری» (bias bias) می‌نامد. تمایل به دیدن هر انحراف از یک معیار آرمانی به‌عنوان خطای

شناختی (یعنی حتی تشخیص سوگیری هم می‌تواند خودش سوگیرانه باشد)، در حالی‌که بسیاری از این الگوها نقص نیستند، بلکه «هیوریستیک‌هایی» (heuristics) هستند سازگار با محیط واقعی خود، و در آن کارآمدند؛ مفهومی که گیگرنزر «عقلانیت بوم‌شناختی» (ecological rationality) می‌نامد.

این نقد تز این مقاله را باطل نمی‌کند، به این دلیل که حتی اگر برخی موارد هیوریستیک سازگارانه باشند نه خطا، باز هم کالایی‌شدن نام آنها بدون فهم شرایط کاربردشان گمراه‌کننده است؛ اما این نقد یادآوری می‌کند که خود مقولهٔ «سوگیری شناختی» یک‌دست و بی‌مناقشه نیست.

اما این تصویر هنوز ناقص است. آنچه تا اینجا دیدیم، نقطهٔ کور سوگیری را عمدتاً در مقام یک محدودیت شناختیِ نسبتاً خنثی توصیف کرد؛ خطایی در خودسنجی که هر فرد، صرف‌نظر از جایگاه اجتماعی یا سیاسی‌اش، کم و بیش به یک اندازه گرفتار آن است. اما پرسش بعدی این است: آیا این نامتقارنی خودش هم، در عمل، به‌طور یکسان میان افراد و گروه‌ها توزیع می‌شود؟ یا این‌که نحوهٔ بروز و کاربرد آن نیز تابع همان شکاف‌های اجتماعی و سیاسی‌ای است که قرار بود صرفاً شناختشان دهد؟

در بخش بعد به این پرسش می‌پردازیم.

---

۲. برچسب سوگیری به‌مثابه سلاح هویتی

اگر نقطهٔ کور صرفاً یک محدودیت شناختی بی‌طرف بود، باید الگوی بروز آن نیز به‌طور یکسان میان همه توزیع می‌شد، اما شاهدی وجود دارد که نشان می‌دهد این نامتقارنی خودش هم سوگیرانه توزیع می‌شود.

واکر، کالینز، واکر، فوگلسانگ و ماندل (Walker, Collins, Walker, Fugelsang & Mandel)، در پژوهشی در دست چاپ (منتشر شده به صورت آنلاین در مه ۲۰۲۶)، در سه آزمایش نشان دادند که هرچه فرد نسبت به شخصی احساس مثبت‌تری داشته باشد، سوگیری کمتری به او نسبت می‌دهد، و این «ارزیابی عاطفی» است که میزان سوگیری را تعیین می‌کند؛ برای نمونه، در آزمایش‌های حزبی، دموکرات‌ها و جمهوری‌خواهان سوگیری به‌مراتب بیشتری به مخالفان حزبی خود نسبت دادند تا به هم‌حزبی‌ها. این یافته، نسخهٔ جدیدتری از چیزی است که ارلینگر، گیلویچ و راس پیشتر در ۲۰۰۵ پنشان داده بودند: «نسبت‌دادن سوگیری به گروه بیرونی» (Ehrlinger, Gilovich & Ross, 2005).

این حلقه پژوهشی را مطالعهٔ کندی و پرونین می‌بندد؛ آنجا که نشان دادند دیدن طرف مقابل به‌عنوان فردی سوگیر، خودش به تشدید منازعه می‌انجامد (Kennedy & Pronin, 2008). این بدان معنی است که برچسب سوگیری نه‌فقط پیامد دشمنی حزبی است، بلکه آن را تقویت هم می‌کند. بنابراین در چنین شرایطی، نسبت‌دادن سوگیری به دیگران، می‌تواند از نقد مستند یک فرآیند استدلالی، به «بی‌اعتبارسازی شخص‌محور» تبدیل شود؛ همچنانکه لی جوسیم، روان‌شناس اجتماعی، در نوشتاری غیرداوری‌شده همین الگو را برای «سوگیری تأیید» توصیف کرده (Jussim, 2019)؛ به‌گفتهٔ او، بدون شواهد مشخص برای گزینش داده یا نادیده‌گرفتن شواهد مخالف، خودِ اتهامِ «سوگیری تأیید» چیزی جز حملهٔ شخصیِ پوشیده نیست.

این کارکرد هویتی، اما، خودش به یک پیش‌شرط نیاز دارد: اینکه نام سوگیری بتواند از پشتوانهٔ دقیق تجربی‌اش جدا شود و به ابزاری ساده، همگان‌فهم و آمادهٔ مصرف در جدل بدل گردد. بخش بعد، نمونه‌ای از همین «فرآیند جدایی» را از نزدیک نشان می‌دهد.

---

۳. اقتصاد توجه: وقتی برچسب از پشتوانهٔ تجربی‌اش جدا می‌شود

همان‌طور که دیدیم، برچسب سوگیری در منازعهٔ هویتی کارکردی پیدا می‌کند که از محتوای تجربی‌اش جدا است؛ و هنگامی که نام یک سوگیری، از زمینهٔ تجربی و شرایط روش‌شناختی خود جدا شود، به کالایی بدل می‌شود که آسان‌تر از خود مهارت تشخیص، پخش می‌شود. نمونهٔ روشن این فرآیند سرنوشت «اثر دانینگ-کروگر» است.

این اثر از نظر فنی، پیش از هر چیز یک پدیدهٔ فراشناختی است (ادراک فرد از توانایی خودش)، نه یک سوگیری ادراکی یا حافظه‌ای از نوع لنگر یا تأیید؛ این تفاوت ماهوی، در کاربرد عمومی و رسانه‌ای نام آن، اغلب نادیده گرفته می‌شود.

ایدهٔ اینکه الگوی دانینگ-کروگر می‌تواند تا حدی محصول برگشت به میانگین و اثر بهتر-از-میانگین در خودارزیابی باشد را نخست کروگر و مولر در سال ۲۰۰۲ به‌طور نظام‌مند مطرح کردند (Krueger & Mueller, 2002). سپس گیگناک و زاژنکوفسکی در ۲۰۲۰، نقدی محدودتر و روش‌شناختی‌تر بر آن افزودند (Gignac & Zajenkowski, 2020)؛ به این ترتیب که از نظر آنها، برای اثبات وجود یک اثر غیرخطی و واقعی، باید جملهٔ درجهٔ دوم معناداری در رگرسیون چند جمله‌ای دیده شود؛ وگرنه الگو می‌تواند صرفاً بازتاب یک رابطهٔ خطی باشد.

این نقدی روش‌شناختی است، نه انکار کامل پدیده؛ همچنانکه یانسن، رفرتی و گریفیتس با رویکرد دیگری استدلال کرده‌اند که بخشی از اثر واقعی باقی می‌ماند (Jansen, Rafferty & Griffiths, 2021).

نکتهٔ مهم برای این مقاله از چگونگی و جزییات این مناقشات، این است که برچسب عمومی‌شده («افراد کم‌توان، مطمئن‌ترند») از دامنهٔ دقیق و محل مناقشهٔ مطالعهٔ اصلی بسیار فراتر رفته، و دقیقاً همین برچسب ساده شده است که در گفتمان سیاسی هم به کار می‌رود: «طرف مقابل نادان است و به همین دلیل مطمئن»؛ الگویی که با کارکرد هویتی برچسب «سوگیری تأیید» در بخش قبل همسو است.

این الگو، یعنی برچسبی که از دامنهٔ دقیق خود فراتر می‌رود و به ابزار جدل بدل می‌شود، ممکن است این برداشت نادرست را القا کند که آموزش دربارهٔ سوگیری‌ها نیز اساساً بی‌فایده است. اما این‌طور نیست؛ بخش بعد نشان می‌دهد آموزشی که به‌درستی طراحی شده باشد، می‌تواند واقعاً اثر داشته باشد؛ هرچند با محدودیت‌های خودش.

---

۴. شواهد مخالف: وقتی آموزش کار می‌کند، و حدودش

ادعای این مقاله محدود است: دانستن نام، به‌تنهایی، ناکافی است؛ نه اینکه هیچ آموزشی اثر ندارد. همانطور که در سال ۲۰۱۵، مورودج، یون، اسکوپلیتی، سیمبورسکی، کوریس و کاسام نشان دادند یک جلسهٔ آموزشی تعاملی سوگیری را تا ۸ هفته (آزمایش ۱) و ۱۲ هفته (آزمایش ۲) بعد، به‌طور قابل‌سنجش کاهش داد (Morewedge et al., 2015)؛ و آموزش ویدیویی غیرتعاملی هم (که تعریف، مثال و راهبرد را توضیح می‌داد، نه صرفاً نام) اثر معنادار و تا حدی پایدار داشت (حدود ۱۹ تا ۲۶ درصد کاهش)، هرچند کمتر از بازی تعاملی. همچنین در آزمایش ۱، خود نقطهٔ کور سوگیری نیز یکی از سه سوگیری‌ای بود که آموزش داده شد و کاهش را نشان داد.

با این توضیح لازم که دو نفر از نویسندگان این مقاله (سیمبورسکی و کوریس)، به Creative Technologies Inc وابسته بودند؛ همان شرکتی که بازی آموزشی «Missing» را ساخته است. این پژوهش در چارچوب برنامهٔ IARPA با نام SIRIUS طراحی شد و با قراردادی از آزمایشگاه تحقیقاتی نیروی هوایی آمریکا تأمین مالی شد. با این حال، مقاله بیانیهٔ رسمی «بدون تضاد منافع» دارد؛ نکته‌ای که صرفاً برای شفافیت یادآوری می‌شود، نه برای تردید در نتایج.

مطالعهٔ میدانی بعدی از این دست، ۲۰۱۹ انجام گرفت و همین اثر آموزشی را در میان ۲۹۰ دانشجوی سه برنامهٔ حرفه‌ای در HEC پاریس، روی کیسی مبتنی بر تصمیم پرتاب شاتل چلنجر، بازتولید کرد (Sellier, Scopelliti & Morewedge, 2019). نتیجه نشان داد که افراد آموزش‌دیده، ۱۹ درصد کمتر از افراد آموزش‌ندیده راه‌حل ضعیف مبتنی بر تأیید فرضیه را انتخاب کردند.

در اینجا برای اینکه بتوانیم این نتایج را به مقیاسی بزرگ‌تر تعمیم دهیم، به دو بازبینی مکمل نیز اشاره می‌کنیم:

نخست، پژوهش سواریاندینی و همکاران، با نوتل (به‌عنوان نویسندهٔ ارشد)، است (Swaryandini et al., 2025). ایشان در ۲۰۲۵، فراتحلیلی از ۴۱ کارآزمایی (۱۶۰ اندازهٔ اثر از میان ۳۸۳ شناسایی‌شده در ۵۴ کارآزمایی، ۱۰٬۹۴۱ شرکت‌کننده) دربارهٔ اثربخشی آموزش‌های کاهش سوگیری در میان دانش‌آموزان و دانشجویان را منتشر کردند. بر اساس نتایج به دست آمده، اثر کلی، بر پایهٔ شاخص اندازهٔ اثر «g هجز» (شاخصی که آماردان لری هجز ابداع کرده و، مانند d کوهن، بزرگی یک اثر را در فراتحلیل می‌سنجد)، کوچک اما معنادار بود (۰.۲۶؛ فاصلهٔ اطمینان ۰.۱۴ تا ۰.۳۹).

گفتنی است که همهٔ این مطالعات «ریسک روش‌شناختی نامشخص یا بالا» داشتند، نشانه‌هایی از سوگیری انتشار هم دیده شد، و نویسندگان نیز صریحاً نوشتند که هنوز معلوم نیست این بهبودها چقدر به تصمیم‌گیری واقعی خارج از کلاس منتقل می‌شوند.

دومین پژوهش، مربوط به فاسولو، هرد و اسکوپلیتی است که در همان ۲۰۲۵، از زاویهٔ دیگری، صد مطالعهٔ تصمیم‌گیری سازمانی را بازبینی کردند (Fasolo, Heard & Scopelliti, 2025). ایشان ۳۲ مطالعه مداخلهٔ آموزشی در سطح فرد را آزموده‌اند، ۶۲ مطالعه طراحی محیط تصمیم را، و ۶ مطالعه هر دو را هم‌زمان. نتیجهٔ آنها این بود که هیچ مطالعه‌ای این دو رویکرد را مستقیماً در برابر هم نسنجیده، و برتری یکی بر دیگری هنوز معلوم نیست.

با این‌همه، به نظر آنها انتخاب میان این دو رویکرد به زمینه بستگی دارد؛ به این معنی که: «اصلاح فردی» (آموزش) برای تصمیم‌های پیچیده و کم‌ساختار، و در شرایطی که زمان و توجه کافی در دسترس است، مناسب‌تر به نظر می‌رسد؛ در حالی‌که «طراحی محیط تصمیم» برای تصمیم‌های ساده‌تر و پایدارتر، و در شرایطی که اهداف افراد همسو است یا نرخ تغییر نیرو بالا است، گزینهٔ بهتری است.

به این ترتیب، این دو بازبینی در کنار هم، تصویر دقیق‌تری به دست می‌دهند که از سه نکتهٔ مستقل تشکیل شده است:

«آموزش سوگیری» واقعاً می‌تواند اثر داشته باشد، اما اندازهٔ اثرش معمولاً کوچک است؛

در مورد «کیفیت روش‌شناختی»، ادبیات این حوزه یکدست نیست، و انتقال آن به تصمیم‌گیری واقعی هنوز پرسشی باز است؛

دربارهٔ «طراحی محیط تصمیم»، به مثابهٔ مسیر مستقل سوم، و در کنار «اصلاح فرد»، نیز باید گفت که هیچ پژوهشی این دو رویکرد را مستقیماً در برابر هم نسنجیده، اما به نظر می‌رسد هرکدام در زمینهٔ خاص خود کارآمدتر است: «آموزش» برای تصمیم‌های پیچیده و کم‌ساختار، در دسترس‌تر و مقیاس‌پذیرتر است، در حالی‌که «طراحی محیط تصمیم» (هرچند نیازمند «معمار تصمیم» و تنظیم جداگانه برای هر زمینه است) برای تصمیم‌های ساده‌تر و پایدارتر مؤثرتر عمل می‌کند.

نتیجهٔ این بخش برای تز مقاله اصلاح‌کننده است، نه نقض‌کننده؛ به این معنی که آنچه کار می‌کند، نه «نام به‌تنهایی»، بلکه «آموزشی با تعریف، مثال، تمرین و بازخورد» است؛ در شرایطی که حتی اندازهٔ اثر این راهکار نیز کوچک و انتقالش به دنیای واقعی هنوز نامطمئن است، و «طراحی محیط تصمیم»، به‌عنوان مسیر مستقل سوم، می‌تواند برای برخی سوگیری‌ها گزینهٔ تکمیلی یا جایگزین باشد.

به این ترتیب، همان «فرانادانی» که این نوشته را با آن آغاز کردیم؛ یعنی ندانستنِ این‌که صرف دانستن نام یک سوگیری، برای اصلاح آن کافی نیست، دست‌کم برای خوانندهٔ این سطور، دیگر «فرانادانی» نیست.

و دست آخر، از مجموع همهٔ این گفته‌ها، شاید بتوان این نتیجهٔ نهایی را گرفت که: نام، آگاهی می‌سازد، ولی صرف دانستن نام یک سوگیری برای اصلاح آن، به همان اندازه کافی است که دانستن نام یک بیماری برای درمان آن: یعنی تقریباً هیچ. سوگیری را نمی‌توان با نامیدنش شکست داد؛ در برابر آن تنها با تمرین و ممارست می‌توان ایستاد.

نتیجه‌گیری

این مقاله، که دومین حلقه از سه‌گانهٔ «توهم کفایت» است، یک ادعای محدود اما مستند را دنبال کرد: دانستن نام یک سوگیری شناختی، شرط لازم اصلاح آن است، نه شرط کافی.

بخش ۱ نشان داد این شکاف صرفاً یک گمان نظری نیست، بلکه در چارچوب ویلسون و برکه قابل صورت‌بندی دقیق است و در آزمایش‌های لنگرسازی و در ادبیات نقطهٔ کور سوگیری، بارها به‌طور رفتاری مستند شده؛ حتی وقتی مکانیسم دقیق نقطهٔ کور همچنان محل مناقشه است، اصل پدیده نیست.

بخش ۲ و ۳ سپس نشان دادند این فرانادانی، در دنیای واقعی، خنثی باقی نمی‌ماند: همان برچسبی که قرار بود ابزار خودشناسی باشد، در منازعات هویتی به سلاحی علیه رقیب بدل می‌شود، و در اقتصاد توجه رسانه‌ها از زمینهٔ دقیق روش‌شناختی‌اش جدا و به شعاری ساده تقلیل می‌یابد؛ سرنوشتی که اثر دانینگ-کروگر نمونهٔ روشن آن است.

بخش ۴ اما نشان داد این تصویر یک‌سویه نیست: وقتی آموزش از صرف دانستن نام فراتر می‌رود و تمرین، بازخورد و کاربرد واقعی را هم در بر می‌گیرد، اثر واقعی و قابل سنجشی بر جای می‌گذارد؛ هرچند آن اثر، طبق شواهد فراتحلیلی موجود، کوچک است، کیفیت روش‌شناختی ادبیاتش یکدست نیست، و انتقالش به تصمیم‌گیری واقعی هنوز نامطمئن. در کنار آموزش، طراحی محیط

تصمیم نیز به‌عنوان مسیری مستقل و گاه مؤثرتر مطرح است، بی‌آنکه تاکنون پژوهشی این دو راهکار را مستقیماً در برابر هم سنجیده باشد.

حاصل جمع این چهار بخش، نه نفی ارزش دانستن نام سوگیری‌ها است و نه تجویز بی قید و شرط آموزش، بلکه ترسیم دقیق مرز میان آنچه «دانستن» می‌تواند انجام دهد و آنچه نمی‌تواند. این مرز، در عین حال، نقطهٔ اتصال این مقاله به بقیهٔ سه‌گانه است؛ به این معنی که: اگر مقالهٔ نخست نشان داد شناخت مغالطات به‌تنهایی کافی نیست، و این مقاله نشان داد شناخت سوگیری‌ها نیز به‌تنهایی کافی نیست، مقالهٔ سوم به شکافی دیگر از همین جنس می‌پردازد؛ شکاف میان داشتن مهارت تفکر انتقادی و گرایش واقعی به کاربردش. فرانادانی، به این ترتیب، نه ویژگی یک حوزهٔ خاص از دانش، بلکه الگویی تکرارشونده در رابطهٔ انسان با دانشی است که دربارهٔ محدودیت‌های خودش دارد.

منابع

Beck, A. T. (1976). Cognitive therapy and the emotional disorders. International Universities Press.

Benson, B. (2016). Cognitive bias cheat sheet. Better Humans (Medium).

Burns, D. D. (1980). Feeling good: The new mood therapy. William Morrow.

Burns, D. D. (1999). Feeling good: The new mood therapy (Rev. ed.). Avon Books.

Chandrashekar, S. P., Yeung, S. K., Yau, K. C., Cheung, C. Y., Agarwal, T. K., Wong, C. Y. J., Pillai, T., Thirlwell, T. N., Leung, W. N., Tse, C., Li, Y. T., Cheng, B. L., Chan, H. Y. C., & Feldman, G. (2021). Agency and self-other asymmetries in perceived bias and shortcomings: Replications of the bias blind spot and extensions linking to free will beliefs. Judgment and Decision Making, 16(6), 1392–1412.

Ehrlinger, J., Gilovich, T., & Ross, L. (2005). Peering into the bias blind spot. Personality and Social Psychology Bulletin, 31(5), 680–692.

Fasolo, B., Heard, C., & Scopelliti, I. (2025). Mitigating cognitive bias to improve organizational decisions. Journal of Management, 51(6), 2182–2211.

Fernbach, P. M., Rogers, T., Fox, C. R., & Sloman, S. A. (2013). Political extremism is supported by an illusion of understanding. Psychological Science, 24(6), 939–946.

Gigerenzer, G. (2018). The bias bias in behavioral economics. Review of Behavioral Economics, 5(3–4), 303–336.

Gignac, G. E., & Zajenkowski, M. (2020). The Dunning-Kruger effect is (mostly) a statistical artefact. Intelligence, 80.

Jansen, R. A., Rafferty, A. N., & Griffiths, T. L. (2021). A rational model of the Dunning-Kruger effect supports insensitivity to evidence in low performers. Nature Human Behaviour, 5, 756–763.

Jussim, L. (2019, August 7). Confirmation bias: Real bias or delegitimization rhetoric? Psychology Today ("Rabble Rouser" blog). https://www.psychologytoday.com/us/blog/rabble-rouser/201908/confirmation-bias-real-bias-or-delegitimization-rhetoric

Kahneman, D. (2011). Thinking, fast and slow. Farrar, Straus and Giroux.

Kennedy, K. A., & Pronin, E. (2008). When disagreement gets ugly: Perceptions of bias and the escalation of conflict. Personality and Social Psychology Bulletin, 34(6), 833–848.

Kristal, A. S., & Santos, L. R. (2021). G.I. Joe phenomena: Understanding the limits of metacognitive awareness on debiasing (Working Paper No. 21-084). Harvard Business School.

Krueger, J., & Mueller, R. A. (2002). Unskilled, unaware, or both? The better-than-average heuristic and statistical regression predict errors in estimates of own performance. Journal of Personality and Social Psychology, 82(2), 180–188.

Lawson, R. (2006). The science of cycology: Failures to understand how everyday objects work. Memory & Cognition, 34(8), 1667–1675.

Mandel, D. R., Collins, R. N., Walker, A. C., Fugelsang, J. A., & Risko, E. F. (2022). Hypothesized drivers of the bias blind spot: Cognitive sophistication, introspection bias, and conversational processes. Judgment and Decision Making, 17(6), 1392–1421.

Manoogian, J., III. (2016). The cognitive bias codex [Infographic].

Morewedge, C. K., Yoon, H., Scopelliti, I., Symborski, C. W., Korris, J. H., & Kassam, K. S. (2015). Debiasing decisions: Improved decision making with a single training intervention. Policy Insights from the Behavioral and Brain Sciences, 2(1), 129–140.

Pronin, E., Gilovich, T., & Ross, L. (2004). Objectivity in the eye of the beholder: Divergent perceptions of bias in self versus others. Psychological Review, 111(3), 781–799.

Pronin, E., & Kugler, M. B. (2007). Valuing thoughts, ignoring behavior: The introspection illusion as a source of the bias blind spot. Journal of Experimental Social Psychology, 43(4), 565–578.

Pronin, E., Lin, D. Y., & Ross, L. (2002). The bias blind spot: Perceptions of bias in self versus others. Personality and Social Psychology Bulletin, 28(3), 369–381.

Rozenblit, L., & Keil, F. (2002). The misunderstood limits of folk science: An illusion of explanatory depth. Cognitive Science, 26(5), 521–562.

Sage, A., Keppens, J., & Yannakoudakis, H. (2025). A survey of cognitive distortion detection and classification in NLP (arXiv:2508.09878). arXiv.

Santos, L. R., & Gendler, T. S. (2014). [Essay on the "G.I. Joe fallacy"]. In J. Brockman (Ed.), This idea must die: Scientific theories that are blocking progress. Harper Perennial. (As cited in Kristal & Santos, 2021.)

Scopelliti, I., Morewedge, C. K., McCormick, E., Min, H. L., Lebrecht, S., & Kassam, K. S. (2015). Bias blind spot: Structure, measurement, and consequences. Management Science, 61(10), 2468–2486.

Seda, L., Martins, I. T. C., Lisboa, T. L. R. C., Batistuzzo, M. C., & Fatori, D. (2024). Theoretical maturation of the "bias blind spot": A preregistered replication study of Pronin, Lin, and Ross (2002) in a Brazilian sample. Collabra: Psychology, 10(1), Article 122158. https://doi.org/10.1525/collabra.122158

Sellier, A.-L., Scopelliti, I., & Morewedge, C. K. (2019). Debiasing training improves decision making in the field. Psychological Science, 30(9), 1371–1379.

Swaryandini, G., Graham, J., Griffith, S., Grilo, V., Ruzzante, F., Zhang, X., Yeung, S. K., Mangiarulo, M., Basarkod, G., Ng, C., Parker, P., Tangen, J., Saeri, A., Grundy, E., Slattery, P., & Noetel, M. (2025). Systematic review and meta-analysis of educational approaches to reduce cognitive biases among students. Nature Human Behaviour, 9(12), 2510–2538.

Walker, A. C., Collins, R. N., Walker, H. E. K., Fugelsang, J. A., & Mandel, D. R. (2026). Everyone I don't like is biased: Affective evaluations and the bias blind spot. Personality and Social Psychology Bulletin. Advance online publication.

West, R. F., Meserve, R. J., & Stanovich, K. E. (2012). Cognitive sophistication does not attenuate the bias blind spot. Journal of Personality and Social Psychology, 103(3), 506–519.

Wilson, T. D., & Brekke, N. (1994). Mental contamination and mental correction: Unwanted influences on judgment and evaluation. Psychological Bulletin, 116(1), 117–142.

Wilson, T. D., Houston, C. E., Etling, K. M., & Brekke, N. (1996). A new look at anchoring effects: Basic anchoring and its antecedents. Journal of Experimental Psychology: General, 125(4), 387–402.
