تحلیل رفتار معتادان در آزمون قمار آیوا با استفاده از مدل یادگیری تقویتی

هدف: هدف این پژوهش، بررسی اختلالات تصمیم‌گیری معتادان به مواد افیونی، در شرایط مخاطره‌آمیز، با استفاده از روش مدل‌سازی شناختی است. **روش:** در این پژوهش، ابتدا نظام تصمیم‌گیری آزمودنی‌ها (گروه شاهد و گروه معتادان درمان‌جو)، با استفاده از نسخهٔ فارسی *آزمون قمار آیوا* بررسی و سپس، یافته‌های آزمون با استفاده از مدل‌سازی شناختی تحلیل شد. در این راستا و بر اساس یادگیری تقویتی و شرایط *آزمون قمار*، مدل‌های مختلف تصمیم‌گیری در شرایط مخاطره‌آمیز *آزمون قمار* طراحی شد. در مرحلهٔ بعد، مدل‌ها در محیط *آزمون قمار* شبیه‌سازی و جداگانه بر داده‌های هر دو گروه برازش داده شدند. سپس، مدلی که مطابقت بیشتری با رفتار آزمودنی‌های آن گروه داشت، مشخص شد. **یافته‌ها:** نتایج نشان داد که هر دو گروه معتادان درمان‌جو (n=۲۱۷) و شاهد (n=۱۳۰) عملکرد ضعیفی در آزمون داشتند (امتیاز خالص کمتر از ۱۰). علاوه بر آن، عملکرد گروه معتاد ضعیف‌تر از گروه شاهد بود ($p<۰/۰۷$). نتایج مدل‌سازی نشان می‌دهد که مدل بهینهٔ هر دو گروه مدلی است که در ارزیابی انتخاب‌های مختلف فقط به تعداد دفعات دریافت پاداش و زیان (و نه مقادیر آنها) توجه می‌کرد. روش تحلیل حساسیت آشکار کرد که مدل بهینه برای گروه معتاد، علاوه بر توجه صرف به تعداد دفعات پاداش و زیان، بیشتر از گروه شاهد تحت تأثیر آسیب‌گریزی قرار دارد. **نتیجه‌گیری:** نتایج فوق گویای آن است که علت کم بودن امتیاز کسب شده در هر دو گروه، نادیده گرفتن اندازهٔ پاداش و زیان و توجه به دفعات آن است. ضمن این که، دلیل پایین بودن بازده گروه معتاد نسبت به گروه شاهد، عدم تعادل آسیب‌گریزی و پاداش‌جویی و گرایش به سمت آسیب‌گریزی است. این گرایش به آسیب‌گریزی می‌تواند بر شرایط خاص معتادان درمان‌جو منطبق باشد.

کلیدواژه‌ها: *یادگیری تقویتی، آزمون قمار، اعتیاد، مدل‌سازی محاسباتی، مدل‌سازی شناختی، نظام تصمیم‌گیری*

امیر دزفولی
دانشکده برق و کامپیوتر، دانشگاه تهران
محمدمهدی کرامتی
دانشکدهٔ مدیریت و اقتصاد، دانشگاه صنعتی شریف
دکتر هومن صفایی
مرکز ملی مطالعات اعتیاد ایران، دانشگاه علوم پزشکی تهران
کارو لوکس
دانشکده برق و کامپیوتر، دانشگاه تهران
دکتر آذرخش مکری
مرکز ملی مطالعات اعتیاد ایران، دانشگاه علوم پزشکی تهران
دکتر حامد اختیاری *
مرکز ملی مطالعات اعتیاد ایران، دانشگاه علوم پزشکی تهران

* نشانی تماس: کارگر جنوبی، جنب بیمارستان فارابی، مرکز ملی مطالعات اعتیاد ایران
Email: h_ekhtiari@razi.tums.ac.ir

Analysis of Behavior of Drug Addicts in Iowa Gambling Task Using Reinforcement Learning Model

Objective: The aim of this study was to assess decision making disorders in opioid dependents in riskful situations, using cognitive modeling. First, the decision making process of participants (treatment-seeking dependents and the normal group) was evaluated using the Farsi version of Iowa gambling task; and then the data were analyzed using cognitive modeling. Different models of decision making in riskful situations in the gambling task were designed based on reinforcement learning and gambling task circumstances. In the next step, the models were simulated in gambling task environment and were separately fitted on the data of both groups. Thereafter, the model with higher correspondence with the subject behavior in each group was determined. **Results:** Both the group of treatment-seeking dependents (n=127) and the control group (n=130) showed a weak performance in the task (pure scores lower than 10). In addition, the performance of the dependent group was weaker than that of the normal group ($p<0.07$). Modeling results indicated that the optimum model in both groups was the model that was focused only on the frequency of reward and punishment (not on their size), in the assessment of different choices. Sensitivity revealing method revealed that the optimum model for the dependent group, in addition to merely focusing on the frequency of reward and punishment, is affected by harm-avoidance more than the control group. **Conclusion:** The above-mentioned results indicate that the reason for the low scores obtained in both groups, is ignoring the size of reward and punishment and focusing on their frequency. Moreover, the reason for the lower output in the dependent group in comparison with the control group is a lack of balance between harm avoidance and reward dependence and a bias toward harm avoidance. The mentioned bias toward harm avoidance could be associated with the specific circumstances of treatment seeking drug dependents.

Keywords: *reinforcement learning; gambling task; drug dependence; computational modeling; decision making process*

Amir Dezfoli
University of Electronic and Computer Science, Tehran University
Mohammad Mehdi Karamati
School of Economics and Management, Sharif University
Houman Safaie
National Center for Addiction Studies, Tehran University of Medical Science
Karo Looks
University of Electronic and Computer Science, Tehran University
Azarakhsh Mokri
National Center for Addiction Studies, Tehran University of Medical Science
Hamed Ekhtiari
National Center for Addiction Studies, Tehran University of Medical Science

Email: h_ekhtiari@razi.tums.ac.ir

مقدمه

اعتیاد را می‌توان رفتار اجباری[1] به سمت مصرف مواد اعتیادآور، علی‌رغم آگاهی فرد معتاد از آثار زیان‌بار آتی آن، معرفی کرد (انجمن روان‌پزشکی آمریکا[2]، ۲۰۰۰). این گونه اختلالات تصمیم‌گیری، ناشی از آثار دراز مدت مداخلات دارو بر نظام تصمیم‌گیری مغز است که عموماً تأثیر آنها به رفتار فرد معتاد در برابر مادهٔ اعتیادآور محدود نمی‌شود و به تصمیم‌گیری‌های روزمرهٔ وی تسری می‌یابد (اختیاری، رضوان‌فر و مکری، ۱۳۸۷).

بر این اساس، آزمایه‌های ارزیابی شناختی[3] سعی می‌کنند، با شبیه‌سازی وجوهی از تصمیم‌گیری در زندگی روزمره، چگونگی تصمیم‌گیریِ افراد را ارزیابی و اندازه‌گیری کنند. هدف از ارزیابی شناختی در درجهٔ اول، سنجش توانایی شناختی افراد و در درجهٔ دوم، کشف علت یک نارسایی شناختی در نمونهٔ مورد مطالعه است.

یکی از متداول‌ترین آزمون‌ها که برای سنجش تصمیم‌گیری افراد در شرایط مخاطره‌آمیز و غیرقطعی به کار می‌رود، *آزمون قمار آیوا*[4] است (بشارا[5]، داماسیو[6]، داماسیو و آندرسون[7]، ۱۹۹۴). این آزمون یک محیط تصمیم‌گیری واقعی برای آزمودنی شبیه‌سازی می‌کند. مطالعات پیشین حاکی از آن است که *آزمون قمار* قادر است گروه افراد دچار سوءمصرف مزمن مواد اعتیادآور را، بر حسب کارآیی تصمیم‌گیری، از افراد سالم تفکیک کند (اختیاری، بهزادی، جنتی و مکری، ۱۳۸۳). به طور کلی، در این آزمون، کارآیی افراد معتاد کمتر از گروه کنترل بود (بشارا و دامسیو، ۲۰۰۲؛ گرنت[8]، کانترجی[9] و لندن[10]، ۲۰۰۰؛ منترسو[11]، هرمان[12]، ناپیر[13]، ابرین[14] و چایلدرس[15]، ۲۰۰۱؛ رترمام-فیولر[16]، شاپتاو[17]، برمان[18] و لندن، ۲۰۰۴؛ استات[19]، راک[20]، کمپبل[21]، بسمیر[22] و فین[23]، ۲۰۰۵).

به طور سنتی، در پژوهش‌های بالینی و ارزیابی شناختی، فرآیند ارزیابی شناختی با در نظر گرفتن یک آزمون و با هدف سنجش یک توانایی شناختی خاص شروع می‌شود. در مرحلهٔ بعد، سعی می‌شود از عملکرد فرد در آزمون، شاخص‌های مناسبی استخراج شود؛ شاخص‌هایی که بر اساس تحلیل‌های آماری، بین مقدار آن شاخص‌ها و عضویت فرد در گروه شاهد یا بیمار همبستگی کافی وجود داشته باشد. در صورت کشف چنین شاخصی، از آن آزمون و شاخص می‌توان به عنوان آزمونی برای ارزیابی توانایی تصمیم‌گیری افراد بهره جست.

این فرآیند گرچه در مورد سازوکار تصمیم‌گیری در گروه بیمار، اطلاعات ارزشمندی فراهم می‌کند، اما به چرایی رفتار مشاهده شده در یک آزمون نمی‌تواند جواب بدهد. مثلاً به این پرسش که چرا امتیاز گروه معتاد در *آزمون قمار* کمتر از گروه شاهد بوده است، نمی‌توان با دسته‌بندی آنها در دو گروه، بر اساس شاخص عملکردشان پاسخ داد. یک روش برای پاسخ دادن به این گونه پرسش‌ها، بهره‌گیری از روش *مدل‌سازی شناختی*[24] است.

مدل‌سازی شناختی (بسمیر و استات، ۲۰۰۲) رویکردی است محاسباتی، برای توصیف فرآیندهای مغزیِ درگیر در یک عمل شناختی. مدل‌سازی شناختی خود گونه‌ای از *مدل‌سازیِ محاسباتی*[25] است که بر این فرض استوار است که هر یک از سازوکارهای عصبی دخیل در یک عمل شناختی، بر داده‌های ورودی حس‌گرها و خروجی‌های بخش‌های دیگر مغز محاسبه‌ای را پیاده‌سازی می‌کنند. مثلاً، در مورد تصمیم‌گیری، این مجموعه از محاسبات، هدایت‌کنندهٔ انتخاب‌های یک تصمیم‌گیرنده است. مدل‌سازی محاسباتی، این مجموعه فعالیت‌ها را به وسیلهٔ روابط

1- compulsive
2- American Psychiatric Association
3- Cognitive Assessment Tasks
4- *Iowa Gambling Task (IGT)*
5- Bechara
6- Damasio
7- Anderson
8- Grant
9- Contoreggi
10- London
11- Monterosso
12- Ehrman
13- Napier
14- O'Brien
15- Childress
16- Rotheram-Fuller
17- Shoptaw
18- Berman
19- Stout
20- Rock
21- Campbell
22- Busemeyer
23- Finn
24- cognitive modeling
25- computational modeling

ریاضی توصیف می‌کند. وجه اصلی تمایز این روش با روش‌های تحلیل آماری آن است که در مدل‌سازی شناختی، یک عمل شناختی بر اساس فرآیندهای شناخته شدهٔ مغزی توضیح داده می‌شود. مثلاً برای توضیح نحوهٔ تصمیم‌گیری، از عناصری مانند حافظه، ارزیابی پاداش و توجه استفاده می‌شود. این ویژگی امکان برقراری ارتباط میان یک رفتار بیمارگونه و زیرساخت‌های روانی و عصبی آن را فراهم می‌کند. تفاوت این روش با الگوی پیشین تحلیل رفتار در یک آزمون شناختی آن است که این روش فقط به استخراج شاخصی از نحوهٔ تصمیم‌گیری فرد بسنده نکرده، کل فرآیند تصمیم‌گیری را مدل‌سازی می‌کند. تفاوت اصلی مدل‌سازی شناختی با روش‌های مدل‌سازی مفهومی این است که در مدل‌سازی شناختی برای توصیف فرآیند، از زبان ریاضیات استفاده می‌شود که این کار، ارزیابی کمّی انطباق یک مدل با واقعیت را ممکن می‌سازد.

در گام اول این پژوهش، عملکرد تصمیم‌گیری دو گروه معتاد و شاهد (به شرحی که در بخش بعد آمده است) به وسیلهٔ *آزمون قمار آیوا* مورد سنجش قرار گرفت. در گام بعد، منشاء رفتار هر دو گروه شاهد و معتاد به کمک مدل‌سازی شناختی رفتار آنها بررسی شد. این مدل‌ها که منطق تصمیم‌گیری آنها در بخش بعد می‌آید، بر اساس *نظریه یادگیری تقویتی*[1] طراحی شده‌اند (ستون[2] و بارتو[3]، ۱۹۹۸).

روش

ابزار پژوهش

آزمون قمار آیوا (ابزاری برای ارزیابی شناختی تصمیم‌گیری): از *آزمون قمار آیوا*، به عنوان چارچوبی برای ارزیابی بسیاری از اختلالات تصمیم‌گیریِ انسان استفاده می‌شود. در نسخهٔ *ABCD* این آزمون که در این پژوهش به‌کار رفته است، چهار دسته کارت ۶۰ تایی پیش‌روی آزمودنی قرار می‌گیرد. آزمودنی در کل ۱۰۰ انتخاب دارد که می‌بایست در هر انتخاب از یکی از چهار دسته، یک کارت بردارد. پس از هر انتخاب، میزان برد یا باخت به او اطلاع داده می‌شود. آزمودنی باید سعی کند انتخاب‌های او بیشترین میزان سود خالص را به دنبال داشته باشد.

کارت‌ها به دو دستهٔ *سودآور* (*C,D*) و *زیان‌آور* (*A,B*) تقسیم می‌شوند. اگر چه میزان سود کارت‌های سودآور کم است، ولی با توجه به کمتر بودن باخت نسبت به سود، در مجموع سودآورند. کارت‌های زیان‌آور نسبت به کارت‌های سودآور، سود بیشتری نصیب عامل می‌کنند، ولی در کل با توجه به بیشتر بودن میزان زیان آنها نسبت به سود، زیان‌آورند. میزان سود (در دسته کارت‌های سودآور) و زیان (در دسته کارت‌های زیان‌آور) به مرور زمان افزایش می‌یابد. گرچه میانگین سود کارت‌های *A* و *B* برابر است، اما تعداد کارت‌هایی که در دستهٔ *A* باخت به همراه دارند بیشتر، ولی میزان باخت هر کدام کمتر است. در مقابل، در کارت‌های دستهٔ *B*، تعداد کارت‌های دارای باخت کمتر، ولی میزان باخت هر کارت بیشتر است. همین قانون در مورد کارت‌های دو دستهٔ *C* و *D* نیز صادق است. در انتهای آزمون، امتیاز خالص آزمودنی به صورت *(C+D)-(A+B)* محاسبه می‌شود (اختیاری و همکاران، ۱۳۸۳).

آزمودنی‌ها

گروه آزمودنی‌های معتاد را ۲۱۷ فرد معتاد به مواد افیونی (بر اساس معیارهای *DSM-IV*) که برای شرکت در برنامه‌های درمانی به مرکز ملی مطالعات اعتیاد ایران مراجعه کرده بودند، تشکیل می‌دادند. گروه شاهد ۱۳۰ نفر از بستگان افراد گروه معتاد بودند که سابقهٔ مصرف مواد (به‌جز سیگار) نداشتند. گروه شاهد از نظر خصوصیات جنسیتی، سن و سطح تحصیلات، هماهنگی قابل‌قبولی با گروه معتادان داشت. خصوصیات جمعیتی این دو گروه در جدول ۱ آمده است.

1- Reinforcement learning theory
2- Sutton
3- Barto

جدول ۱- خصوصیات جمعیت‌نگاشتی آزمودنی‌های *آزمون قمار*

	گروه‌ها	شاهد	معتاد
شاخص‌های جمعیت‌نگاشتی	تعداد	۱۳۰ نفر	۲۱۷ نفر
	سن (سال)	۳۰/۲۵ ± ۸/۷۷	۲۹/۸۷ ± ۷/۶۰
	تحصیلات (سال)	۱۲/۰۳ ± ۳/۷۷	۱۱/۴۹ ± ۳/۰۰

مدل‌های تصمیم‌گیری

در این بخش از مقاله، دربارهٔ مدل‌های مختلف تصمیم‌گیری مورد استفاده در این پژوهش که در کل ۱۲ مدل خانواده است، توضیح داده می‌شود. تمام مدل‌ها از سه بخش اصلی تشکیل شده‌اند (شکل ۱).

بخش اول در همهٔ مدل‌ها مشترک و مسؤول ارزیابی پاداش و ضرر دریافتی پس از انتخاب یک کارت است.آزمودنی پس از انتخاب کارت (هریک از کارت‌های A, B, C یا D)، پاداش مثبت $r_t^+(a)$ یا پاداش منفی $r_t^-(a)$ ، یا هر دو را دریافت می‌کند. مثلاً، اگر وی پس از برداشتن یک کارت از دسته کارت A، اعداد ۱۰۰+ (میزان برد) و ۲۰۰- (میزان باخت) را در پشت کارت مشاهده کند، ۱۰۰+ همان $r_t^+(a)$ و ۲۰۰ همان $r_t^-(a)$ است. در بخش ارزیابی پاداش، پاداش‌های مثبت و منفی دریافت‌شده به صورت خطی و به وسیلهٔ پارامتری به نام پارامتر ارزش‌گذاری[1] که آن را با w نشان می‌دهیم، با یکدیگر ترکیب شده و از حاصل آن در مرحلهٔ بعد استفاده می‌شود. پارامتر w که مقدار آن در بازهٔ صفر و یک قرار دارد، چگونگی این ترکیب را تعیین می‌کند. مقادیر w نزدیک به یک بیانگر این است که در یادگیری، مدل فقط مقادیر پاداش‌های مثبت دریافت شده ($r_t^+(a)$) را در نظر می‌گیرد و به مقادیر ضررها ($r_t^-(a)$) بی‌اعتناست. از طرف دیگر، مقادیر نزدیک به صفر بدان معناست که عامل فقط به ضررهای دریافتی توجه می‌کند و به پاداش‌های مثبتِ دریافتی بی‌توجه است. در اینجا ما حاصل ترکیب پاداش‌های مثبت و منفی را پاداش دریافت شده می‌نامیم. به این گونه یادگیری که فقط بر اساس مشاهدهٔ پاداش و ضرر انجام می‌شود، یادگیری تقویتی می‌گویند. در واقع منطق زیر بنایی این نوع یادگیری این است که آزمودنی فقط بر اساس مشاهدهٔ پاداش و ضرر تصمیم‌گیری می‌کند و منبع اطلاعاتی دیگری برای تصمیم‌گیری در اختیار ندارد. به نظر می‌رسد که این الگوی تصمیم‌گیری با رفتار آزمودنی‌ها در *آزمون قمار آیوا* انطباق قابل قبولی داشته باشد، زیرا منبع مشاهدهٔ آزمودنی در این آزمون هم فقط پاداش‌ها و ضررهای دریافت شده است.[1]

بخش دوم، بخش یادگیرنده است که از پاداش‌های دریافت شده برای یادگیری استفاده می‌کند. خروجی مرحلهٔ یادگیری، تخمین ارزش هر دسته از کارت‌هاست که این تخمین‌ها در شکل ۱ به وسیلهٔ Q(A)، Q(B)، Q(C) و Q(D) نشان داده شده‌اند. مثلاً مقدار Q(D) در سی‌امین انتخاب مدل، نشان‌دهندهٔ این است که مدل با انتخاب کارت از دستهٔ D، انتظار دارد چه عایدش بشود. به این نوع یادگیری که به انتخاب‌های مختلف ارزش می‌دهد، *یادگیری مبتنی بر ارزش*[2] گفته می‌شود. بر خلاف بخش پیشین که در همهٔ مدل‌های استفاده شده مشابه است، این بخش در مدل‌های مختلف متفاوت است و هر مدل برای یادگیری ارزش هر دسته کارت، از روش‌های مختلف استفاده می‌کند.

الگوهای ارائه‌شده، با مطالعهٔ یافته‌های آزمون قمار در بیماران دچار ضایعات مغزی در ناحیهٔ میانی تحتانی قشر پیش‌پیشانی[3] به دست آمده است (کالیدیندی[4]، باومان[5]، ۲۰۰۷). در کل از شش الگوی مختلف یادگیری استفاده شده است که می‌توان آنها را به دو دسته تقسیم کرد. دستهٔ اول مدل‌ها برای یادگیری ارزش هر دسته از کارت‌ها، از مقدار پاداش دریافت‌شده استفاده می‌کند. در یکی از مدل‌های این دسته، برای پیش‌بینی ارزش هر دسته از پاداش‌های دریافت‌شده قبلی، میانگین گیری می‌شود. در این صورت، مقدار ارزش دسته کارت D مثلاً در انتخاب سی‌ام برابر است با میانگین پاداش‌های کسب شده در انتخاب‌های قبلی از این دسته کارت. این دسته از مدل‌ها را *یادگیری بر اساس مقدار* می‌نامیم. تفاوت مدل‌های دسته اول و دوم این است که این مدل‌ها به مقدار پاداش دریافت شده بی‌توجه‌اند و فقط به این توجه دارند که حاصل پاداش دریافت شده مثبت است یا منفی. برای مثال، این که پس از انتخاب یک

1- valence parameter
2- value-based learning
3- ventro medial prefrontal cortex
4- Kalidindi
5- Bowman

کارت، پاداش دریافت شده ۱۰۰+ یا ۵۰۰+ باشد، بر پیش‌بینی‌های بعدی در مورد ارزش این دسته کارت تأثیر یکسانی دارند، ولی اگر مثلاً مقدار پاداش دریافت شده ۵۰- باشد (باخت یا تنبیه)، تأثیر متفاوتی بر پیش‌بینی‌های بعدی گذاشته و مقدار ارزش دسته کارت را کم می‌کند، اگرچه مقدار کم شدن این ارزش در صورتی که مدل پاداش ۶۰۰- را هم دریافت می‌کرد، با وضعیتی که ۵۰- دریافت کرده بود، تفاوتی نمی‌کرد. ما الگوهای این دسته را *یادگیری مبتنی بر تعداد* می‌نامیم.

روشن است که نادیده گرفتن مقدار پاداش و یادگیری، صرفاً بر اساس مثبت یا منفی بودن پاداش دریافت شده، عامل را در تخمین ارزش کارت‌ها به اشتباه می‌اندازد. اما هدف از در نظر گرفتن مدل‌های مختلف این است که بررسی شود کدام یک از مدل‌ها بر رفتار مشاهده شده در آزمون بیشتر منطبق است. مثلاً، اگر مدل‌های دستهٔ دوم بهتر از دستهٔ اول بر رفتار یک فرد در *آزمون قمار* برازش پیدا کرد، می‌توان نتیجه گرفت که فرد بیش از آن که تحت تأثیر مقدار پاداش و ضرر باشد، متأثر از دفعات آنهاست. هر دو دسته مدل‌های یادگیری انواع مختلف دارند که از شرح آنها و روابط ریاضی هر یک صرف‌نظر کرده و خوانندهٔ محترم را به ضمیمه ارجاع می‌دهیم.

چگونگی یادگیری در الگوهای یادگیری به وسیلهٔ پارامتر *نرخ یادگیری* (که آن را با γ نشان می‌دهیم) کنترل می‌شود. این پارامتر در بازهٔ صفر و یک قرار دارد و اگر مقدار آن در یک مدل نزدیک به یک باشد، بیانگر آن است که مدل فقط به تجارب جدید خود برای تخمین ارزش هر دسته کارت توجه می‌کند و به تجارب قبلی اهمیت نمی‌دهد. از طرف دیگر، مقادیر نزدیک به صفر نشان می‌دهد که اثر تجارب جدید بر یادگیری ارزش هر دسته از کارت‌ها کم است و تأثیر تجارب جدید بر ارزش‌ها به‌کندی صورت می‌گیرد.

مرحلهٔ سوم، انتخاب یک کارت بر اساس تخمین ارزش هر دسته کارت است. در واقع، عامل، بر اساس ارزش‌هایی که در مرحلهٔ قبل به هر انتخاب خود نسبت داده است، یک کارت انتخاب می‌کند. به این روش تصمیم‌گیری که در آن عامل به گزینه‌های مختلف پیش‌رو ارزش نسبت می‌دهد و سپس بر اساس ارزش نسبی آنها تصمیم‌گیری می‌کند، تصمیم‌گیری مبتنی بر ارزش گفته می‌شود. تصمیم‌گیری مبتنی بر ارزش صرفاً الگوی تصمیم‌گیری در شرایط همراه با عدم قطعیت و مخاطره‌آمیز نیست، گرچه از پشتوانهٔ عصبی محکمی برخوردار است و الگویی منسجم برای تصمیم‌گیری در حین یادگیری ارائه می‌کند؛ به همین علت نیز در مطالعهٔ حاضر از این الگوی تصمیم‌گیری استفاده شده است.

در این پژوهش دو الگوی انتخاب عمل مختلف به کار رفت. در الگوی اول، که روش ε- *حریصانه*[1] نام دارد، مدل با احتمال ε یک دسته کارت را به تصادف انتخاب می‌کند و با احتمال ε-۱ انتخابش از دسته کارتی است که پیش‌بینی می‌کند بیشترین پاداش با انتخاب آن کارت عایدش می‌شود. اگر عامل همیشه کارت را از دسته‌ای انتخاب کند که بیشترین ارزش تخمین‌زده شده را دارد، ممکن است قبل از این که به تخمین صحیحی از ارزش هر دسته کارت برسد، تحت تأثیر انتخاب‌های اولیه به یک دسته کارت گرایش پیدا کند. انتخاب‌های تصادفی به منظور پاسخ‌گویی به این مشکل است، زیرا باعث می‌شود مدل تعدادی از انتخاب‌های خود را به اکتشاف محیط اختصاص دهد تا قبل از گرایش به یک دسته کارت، به تخمین درستی از ارزش آنها برسد. مقادیر نزدیک به یکِ ε، مدل را به سمت رفتار اکتشافی زیاد هدایت می‌کند. از طرفی، مقادیر نزدیک به صفر باعث می‌شود مدل از اکتشاف خودداری کرده و بر اساس تخمین‌های قبلی خود از ارزش هر کارت، انتخاب عمل کند.

در الگوی دوم که *سافت- مَکس*[2] نام دارد، مدل از هر دسته کارت با احتمالی متناسب با ارزش آن دسته کارت، کارت انتخاب می‌کند. بدین معنا که اگر دسته کارتی ارزش تخمین زده شده زیادی داشته باشد، احتمال انتخاب آن دسته نسبت به بقیه دسته‌ها به وسیلهٔ عامل بیشتر می‌شود. این که چقدر زیاد بودن ارزش هر دسته کارت منجر به انتخاب آن با احتمال بیشتر

1- greedy 2- soft-max

می‌شود، با پارامتر میزان اکتشاف- بهره‌برداری[1] که آن را با β نشان می‌دهیم، کنترل می‌شود. هرچه مقدار این پارامتر بزرگ‌تر باشد، احتمال انتخاب دسته کارت با ارزش بالاتر بیشتر می‌شود و مدل رفتار اکتشافی کمتری دارد.

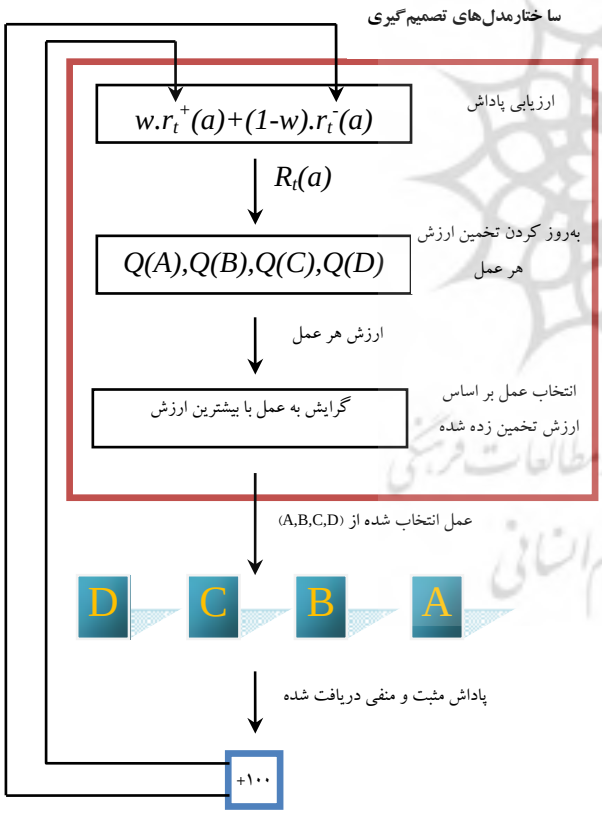
از ترکیب شش الگوی یادگیری با دو الگوی انتخاب عمل، ۱۲ مدل به دست می‌آید که در این مطالعه، هرکدام جداگانه بر رفتار مشاهده از هر گروه برازش داده شدند و کیفیت برازش آنها با یکدیگر مقایسه شد. برای کشف این که کدام مدل بیشتر بر رفتار یک گروه منطبق بوده است، از معیار اطلاعات بیزی[2] استفاده شد که شرح آن در بخش ضمیمه آورده شده است. برای هر گروه به طور جداگانه مقادیر پارامترهای آزاد مدل (مانند w، γ) تخمین زده شد. منظور از تخمین پارامتر، شناسایی مقداری از پارامتراست که وقتی مدل با آن مقدار شبیه‌سازی می‌شود، رفتاری نزدیک‌تر به رفتار مشاهده شده از گروه مورد نظر تولید می‌شود.

یافته‌ها

نتایج *آزمون قمار* در دو گروه شاهد و معتاد در جدول ۲ آمده است. این نتایج در واقع اختلاف تعداد انتخاب‌ها از دسته کارت‌های سودآور (C+D) و دسته کارت‌های زیان‌آور (A+B) می‌باشد. همان‌طور که دیده می‌شود، میانگین امتیاز کسب شده در گروه شاهد ۶/۸۸ و بیشتر از گروه معتاد (۲/۳-) است که این اختلاف به معنادار بودن نزدیک است (p=۰/۰۷). با وجود این، امتیاز گروه شاهد در مقایسه با نمونه‌های غربی کم است. در واقع، مطالعات غربی حداقل امتیاز ۱۰ را معرف عملکرد سالم گروه شاهد می‌دانند (بشارا و همکاران، ۲۰۰۱).

در سطر دوم جدول، اختلاف انتخاب‌ها از دسته کارت‌های با فرکانس پاداش بالا (B+D) و دسته کارت‌های با فرکانس پاداش پایین (A+C) نشان داده شده است. همان‌طور که دیده می‌شود، آزمودنی‌های هر دو گروه گرایش معناداری به دسته کارت‌های با فرکانس پاداش بالا داشتند. در ادامهٔ این بخش علت اختلاف امتیاز دو گروه شاهد و معتاد، و همچنین آزمودنی‌های شاهد ایرانی و مطالعات غربی بررسی شده است. همچنین، علت گرایش هر دو گروه به دسته کارت‌های با فرکانس پاداش بالا بررسی می‌شود.

از میان انواع روش‌های یادگیری تقویتی، مدل بهینه برای هر دو گروه، مدل یادگیری مبتنی بر تعداد به همراه روش انتخاب عمل سافت- مکس است. همان‌طور که در بخش قبل گفته شد، در مدل یادگیری مبتنی بر تعداد، مقادیر برد و باخت در یادگیری نقشی ندارند و به جای آنها از تعداد برد و باخت برای یادگیری استفاده می‌شود. بهتر بودن توانایی این مدل در توصیف رفتار هر دو گروه بیان‌کنندهٔ آن است که کم بودن امتیاز خالص هر دو گروه به علت نادیده‌گرفتن مقادیر برد و باخت و توجه به تعداد برد و باخت و در نتیجه انتخاب از کارت‌های زیان‌آور بوده است.



شکل ۱- ساختار مدل‌های تصمیم‌گیری مرتبط با *آزمون‌های قمار*

1- Exploration-exploitation balance
2- Bayesian information criterion

جدول ۲- *نتایج آزمون قمار* بر اساس جمع و تفریق دفعات انتخاب از کارت‌های C, B, A یا D در طی آزمون (مجموع ۱۰۰ انتخاب)

	گروه‌ها	شاهد	معتاد
نمرات آزمون قمار	(C+D)-(A+B)	۶/۸۸ ± ۲۸/۴۷	-۲/۳۹ ± ۲۶/۱۰
	(B+D)-(A+C)	۱۹/۶۵± ۲۱/۵۵	۱۷/۵۲ ± ۲۱/۸۵

مقادیر بهینهٔ پارامترها به ازای این مدل به‌صورت زیر است:

$$P^*_{control} = (\hat{\beta} = ۰/۷۰ \ , \hat{w} = ۰/۳۵ \ , \hat{\gamma} = ۰/۲۰ \ )$$
$$P^*_{addict} = (\hat{\beta} = ۰/۵۰ \ , \hat{w} = ۰/۱۵ \ , \hat{\gamma} = ۰/۳۵ \ ) \qquad (۱)$$

که در آن، $\hat{\beta}$ میزان اکتشاف- بهره‌برداری، $\hat{w}$ پارامتر ارزش‌گذاری و $\hat{\gamma}$ نرخ یادگیری است. بر اساس مدل بهینه برای هر گروه می‌توان گفت:

از آنجا که مدل یادگیری بر اساس تعداد دفعات، نسبت به مدل یادگیری بر اساس مقدار، توصیف‌کنندهٔ بهتری برای رفتار هر دو گروه است، می‌توان نتیجه گرفت که هر دو گروه بیش از آن که به مقادیر پاداش و تنبیه (برد و باخت) حساس باشند، به تعداد دفعات آنها حساسیت دارند. (شکل ۲)

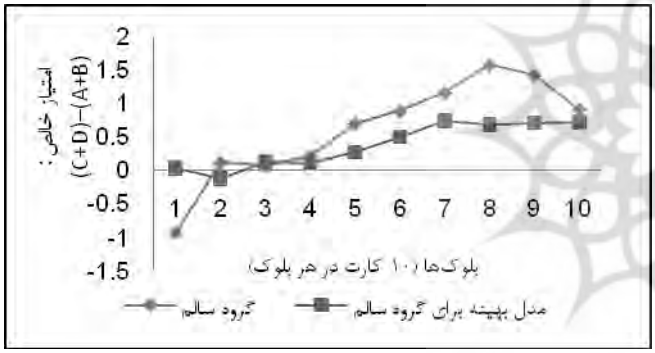
۲. زیادتر بودن مقدار پارامتر β در گروه شاهد نشان می‌دهد که گروه شاهد بیشتر از گروه معتاد، تمایل به انتخاب اعمال با ارزش انتظاری بیشتر دارد.

۳. با توجه به بزرگ‌تر بودن مقدار پارامتر w در گروه معتاد، می‌توان نتیجه گرفت که رفتار آسیب‌گریزی در گروه معتاد، بیش از گروه شاهد است.

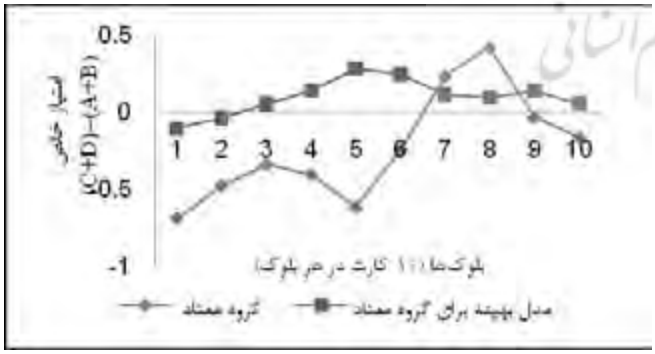
۴. بزرگ‌تر بودن نرخ یادگیری در گروه شاهد حاکی از آن است که این گروه به تجارب جدید اهمیت بیشتری می‌دهد.

متوسط نتایج آزمون قمار و کارآیی مدل بهینه در گروه شاهد در شکل ۳ و در گروه معتاد در شکل ۴ مقایسه شده است.

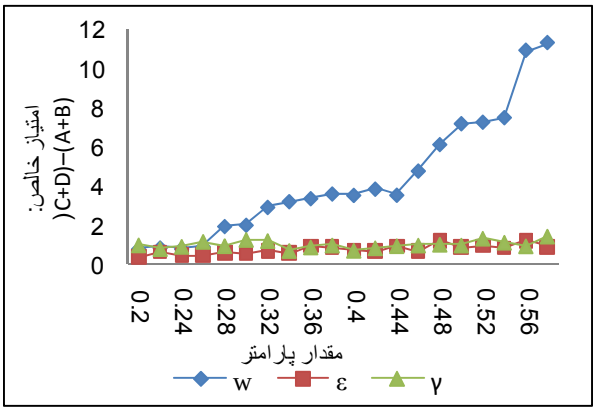
شکل ۲- عامل متأثر از دفعات ضرر و پاداش. ارزش دریافت سه بار پاداش ۱۰۰ واحدی برای عامل، بیشتر از دریافت یک پاداش ۷۰۰ واحدی است.



شکل ۳- کارآیی متوسط گروه شاهد و مدل بهینه برای گروه شاهد



شکل ۴-کارآیی متوسط گروه معتاد و مدل بهینه برای گروه معتاد



شکل ۵- کارآیی مدل به ازای مقادیر مختلف پارامترها در اطراف بردار P^{*}_{addict}

شکل ۶- عامل با گرایش به آسیب‌گریزی. تحلیل رفتار معتادان در آزمون قمار ایوا با استفاده از مدل یادگیری تقویتی. برای یک عامل دریافت یک پاداش ۵۰ واحدی ارزشمندتر از یک پاداش ۱۰۰ تومانی و یک حذر ۵۰ تومانی است.

با توجه به این که هر دو گروه شاهد و معتاد، تعداد پاداش و تنبیه را بیش از مقادیر آنها در تصمیم‌گیری لحاظ می‌کنند، باید به دنبال عامل متمایزکنندهٔ دیگری گشت که بتواند اختلاف در سطح کارآیی دو گروه را توضیح دهد. شکل ۵، کارآییِ مدل یادگیری مبتنی بر تعداد و روش انتخاب عمل سافت- مکس را به ازای انحراف هر یک از پارامترها از نقطهٔ بهینهٔ P^{*}_{addict} نشان می‌دهد.

همان‌طور که شکل ۵ نشان می‌دهد، کارآیی مدل حساسیت زیادی به تغییرات پارامتر ارزش‌گذاری (w) دارد و هنگامی که مقدار این پارامتر به مقدار متناظر خود در مدل بهینهٔ گروه شاهد می‌رسد، کارآیی مدل بهینهٔ گروه معتاد به کارآیی مدل بهینهٔ گروه شاهد می‌رسد. بنابراین، می‌توان نتیجه‌گرفت که انحراف افراد معتاد به رفتار آسیب‌گریزی در کنار تأثیرپذیری زیاد از تعداد دفعات پاداش (به‌جای مقادیر آن) عوامل توضیح دهندهٔ کارآیی پایین گروه معتاد هستند. (شکل ۶)

نتیجه‌گیری

در این پژوهش از دو گروه معتاد و شاهد با استفاده از *آزمون قمار آیوا* ارزیابی شناختی به عمل آمد. نتایج حاصل حاکی از ضعف تصمیم‌گیری هر دو گروه در *آزمون قمار* و ضعیف‌تر بودن گروه معتاد نسبت به گروه شاهد در این آزمون است. همان‌طور که در بخش مقدمه ذکر شد، در مطالعات پیشین نیز ناتوانی افراد وابسته به مواد اعتیادآور در *آزمون قمار* مشاهده شده بود.

برای بررسی علت این مشاهدات از روش مدل‌سازی شناختی استفاده شد. بدین منظور، انواع یادگیری تقویتی بر رفتار هر دو گروه برازش داده شدند و توانایی توصیف هر الگو موردبررسی قرار گرفت. نتایج حاکی ازآن بود که هر دو گروه متأثر از تعداد دفعات برد و باخت بودند، اما به مقدار آن توجه نداشتند. این عامل می‌تواند ضعف هر دو گروه را در آزمون توضیح دهد. از طرف دیگر، تحلیل حساسیت مدل بهینه برای گروه معتاد، این واقعیت را آشکار کرد که گرایش افراد معتاد به آسیب‌گریزی بیشتر از افراد سالم است که این عامل عملکرد ضعیف‌تر افراد معتاد را (نسبت به افراد سالم) در آزمون تبیین می‌کند.

در مطالعات پیشین، برای تحلیل علت ناکارآمدی معتادان در *آزمون قمار* از روش مدل‌سازی شناختی استفاده شده است (استات، بسمیر، لین[1]، گرنت و بنسن[2]، ۲۰۰۴؛ استات و همکاران، ۲۰۰۵).

1- Lin 2- Bonson

در این مطالعات فقط مدل مبتنی بر مقدار پاداش بر رفتار معتادان برازش داده شدند (اما مدل مبتنی بر دفعات بر رفتار گروه‌ها برازش داده نشدند) که نتایج آنها نشان می‌دهد گرایش معتادان به پاداش‌جویی علت اصلی تفاوت رفتار آنها با گروه شاهد است. در واقع این مطالعات عملکرد ضعیف معتادان را، نه در کاستی حافظه یا یادگیری، بلکه ناشی از عدم تعادل در توجه به پاداش و ضرر یافته‌اند. از این منظر، یافتهٔ این مطالعات بر یافته‌های این پژوهش منطبق است.

علت اصلی تفاوت عملکرد گروه معتاد با گروه شاهد در پژوهش حاضر نیز تفاوت آنها در میزان توجه به پاداش و ضرر بوده است. از نظر انتخاب عمل، در مطالعات پیشین نیز (مانند این پژوهش) انتخاب عملِ معتادان کمتر از گروه شاهد تحت تأثیر ارزش پیش‌بینی شده برای هر عمل قرار داشت (استات و همکاران، ۲۰۰۴). تنها تفاوت یافته‌های این مطالعه با مطالعات پیشین این است که علت کارآیی پایین معتادان در این مطالعه گرایش به آسیب‌گریزی اعلام شده است، در حالی که مطالعات قبلی، عدم آسیب‌گریزی و گرایش به پاداش‌جویی را علت اصلی تشخیص داده بودند.

دلایل تفاوت یافته‌های این پژوهش با مطالعات پیشین را می‌توان در موارد زیر خلاصه کرد:

مطالعات پیشین (نسبت به مطالعهٔ فعلی) برای مدل‌سازی رفتاری از مدل‌های محدودتری استفاده کرده بودند که شامل مدل یادگیری مبتنی بر دفعات پاداش نبود و از آنجا که مدل بهینهٔ استفاده شده در این پژوهش، مدل یادگیری مبتنی بر تعداد دفعات بوده (و نه مدل یادگیری مبتنی بر مقدار پاداش)، نتایج این مطالعات را نمی‌توان مستقیماً با یکدیگر مقایسه کرد. به همین دلیل نیز، معنای گرایش به آسیب‌گریزی در مدل مورد استفادهٔ این پژوهش با معنای گرایش به آسیب‌گریزی در مدل یادگیری مبتنی بر مقدار تفاوت دارد. بنابراین، برای مقایسهٔ دقیق‌تر باید مدل یادگیری مبتنی بر دفعات را بر داده‌های استفاده شده در پژوهش‌های پیشین نیز برازش داده و آن را با مدل یادگیری مبتنی بر مقدار مقایسه کرد.

علت دیگر متفاوت بودن نتایج این مطالعه با مطالعات پیشین را می‌توان به درمان‌جو بودن معتادان این مطالعه نسبت داد. معتادانی که در این پژوهش مورد ارزیابی شناختی قرار گرفتند، افرادی بودند که برای درمان به مرکز ملی مطالعات اعتیاد مراجعه کرده بودند. اساساً معتادانی که برای درمان مراجعه می‌کنند، افرادی هستند که متأثر از تجارب گذشتهٔ خود، در پی جلوگیری از آسیب‌های آتی سوء مصرف مواد برآمده و برای درمان مراجعه کرده‌اند. بنابراین، رفتار آسیب‌گریزی مشاهده شده در آزمون را می‌توان با میل به دوری از آسیب سوء مصرف مواد اعتیادآور (که به خصلت آسیب‌گریزی این گروه منجر شده است) توجیه کرد. در مطالعات پیشین، گروه معتاد، معتادان درمان‌جو نبودند و بنابراین می‌توان از آنها رفتاری متفاوت از رفتار معتادان درمان‌جوی این پژوهش انتظار داشت.

به لحاظ عصبی تصمیم‌گیری بر اساس تعداد (به جای مقدار پاداش) در گروه معتاد را می‌توان به کاهش پردازش‌های عصبی در ناحیهٔ پیش‌پیشانی نسبت داد که بر اثر مصرف مزمن مواد مخدر رخ داده است. این کاهش در سطح فعالیت بخش قشری پیشانی[1] می‌تواند فعالیت نواحی دوپامینیِ تحت قشری را افزایش دهد. بر اثر این دو رویداد، سازوکار تصمیم‌گیری تا حد زیادی به هسته‌های قاعده‌ای[2] سپرده می‌شود که، بر اساس مطالعات صورت‌گرفته، در بازنمایی تعداد پاداش و تنبیه نقشی اساسی دارند (فرانک[3] و کلاوس[4]، ۲۰۰۶).

علت کارآیی کمتر گروه سالم این پژوهش از گروه سالم مطالعات غربی می‌تواند تفاوت‌های بین فرهنگی باشد. رواجِ کم قمار در ایران به علت اعتقادات دینی، موجب کاهش آشناییِ عمومی با این مفهوم شده است. این امر به نوبهٔ خود می‌تواند به کاهش کارآیی آزمودنی‌های ایرانی در *آزمون قمار* بینجامد.

در ادامهٔ پژوهش ارائه‌شده در این مقاله، می‌توان به تحلیل نظام تصمیم‌گیری معتادان غیردرمان‌جو پرداخت. این کار

1- frontal cortical
2- basal ganglia
3- Frank
4- Claus

می‌تواند به روشن شدن علت تفاوت نتایج این پژوهش با نمونه‌های خارجی کمک کند.

این نکته نیز اهمیت دارد که در این پژوهش، میزان خوب بودن برازش مدل بهینه هر گروه به داده‌های آن گروه، به نحو مطلوب محاسبه نشده است که علت آن دشواری این کار در مدل‌های غیر خطی (مانند مدل‌های استفاده شده در این پژوهش) است . برای محاسبهٔ مطلوب میزان خوب بودن برازش مدل بهینهٔ هر گروه به داده‌های آن گروه، می‌توان به توسعهٔ روابط ریاضی در مطالعات بعدی توجه کرد.

برای تأیید این فرض که علت کاستی در تصمیم‌گیری، توجه به تعداد دفعات پاداش و نه مقدار آنها بوده است، لازم است آزمونی طراحی شود که در آن توجه به دفعات در تصمیم‌گیری به کارآیی بالا در آزمون منجر شود. در این صورت اگر گروه سالم در آزمون کارآیی بالایی داشته باشد، می‌توان نتیجه گرفت که مبنای تصمیم‌گیری آزمودنی‌های ایرانی تعداد دفعات معیار بوده و کم بودن امتیاز در این پژوهش به علت کم‌توجهی آزمودنی‌ها به آزمون یا ناآشنایی با *آزمون قمار* نبوده است.

برای تأیید دستاوردهای این پژوهش می‌توان در پژوهش‌های بعدی، از روایی‌سنجی خارجی[1] استفاده کرد؛ بدین معنا که مثلاً ارتباط بین میزان انحراف به سمت آسیب‌گریزی و *شاخص شدت اعتیاد*[2] بررسی یا در مورد ارتباط میان هوش‌بهر و میزان گرایش به تصمیم‌گیری بر اساس تعداد دفعات مطالعه شود.

در نهایت نیز به منظور روایی‌سنجی مدل برازش‌یافته بر رفتار یک گروه، استفاده از ثبت فعالیت‌های مغزی در حین انجام آزمون شدنی است؛ بدین صورت که ارتباط میان سیگنال‌های درونی مدل و فعالیت بخش‌های مختلف مغز در حین آزمون بررسی شود. در صورت کشف تناظر قابل توجه می‌توان نتیجه گرفت که در واقع محاسبهٔ انجام شده با مدل، در نظام عصبی پیاده‌سازی می‌شود (کرادو[3] و دیا[4]، ۲۰۰۷).

دریافت مقاله: ۱۳۸۸/۱/۱۶؛ پذیرش مقاله: ۱۳۸۹/۲/۹

1- external validation
2- *Addiction Severity Index*
3- Corrado
4- Doya

منابع

اختیاری، ح.، بهزادی، آ.، جنتی، آ.، و مکری، آ. (۱۳۸۳). دفعات باخت و مقادیر آن: کدام یک تأثیر منفی بیشتری بر ما می‌گذارد؟ *فصلنامهٔ تازه‌های علوم شناختی، ۳ و ۴*، ۲۷-۱۷.

اختیاری، ح.، رضوانفر، م.، و مکری، آ. (۱۳۸۷). تکانشگری و ابزارهای گوناگون ارزیابی آن: بازبینی دیدگاه‌ها و بررسی‌های انجام شده. *مجله روان‌پزشکی و روان‌شناسی بالینی ایران*، ۴ (۳)، ۲۵۷-۲۴۷.

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders: DSM-IV-TR (Text Revision)*. Washington: American Psychiatric Association.

Bechara, A., Damasio, A. R., Damasio, H., & Anderson, S. W. (1994). Insensitivity to future consequences following damage to human prefrontal cortex. *Cognition*, *50*(1-3), 7-15.

Bechara, A., Dolan, S., Denburg, N., Hindes, A., Anderson, S. W., & Nathan, P. E. (2001). Decision-making deficits, linked to a dysfunctional ventromedial prefrontal cortex, revealed in alcohol and stimulant abusers. *Neuropsychologia*, *39*(4), 376-389.

Bechara, A., & Damasio, H. (2002). Decision-making and addiction (part I): Impaired activation of somatic states in substance dependent individuals when pondering decisions with negative future consequences. *Neuropsychologia*, *40*(10), 1675-1689.

Busemeyer, J. R., & Stout, J. C. (2002). A contribution of cognitive decision models to clinical assessment:

Decomposing performance on the Bechara gambling task. *Psychological Assessment, 14*(3), 253-262.

Corrado, G., & Doya, K. (2007). Understanding neural coding through the model-based analysis of decision making. *Journal of Neuroscience, 27*(31), 8178-8180.

Frank, M. J., & Claus, E. D. (2006). Anatomy of a decision: Striato-orbito frontal interactions in reinforcement learning, decision making, and reversal. *Psychological Review, 113*(2), 300-326.

Grant, S., Contoreggi, C., & London, E. D. (2000). Drug abusers show impaired performance in a laboratory test of decision making. *Neuropsychologia, 38*(8), 1180-1187.

Kalidindi, K., & Bowman, H. (2007). Using e-greedy reinforcement learning methods to further understand ventromedial prefrontal patients' deficits on the Iowa Gambling Task. *Neural Networks, 20*(6), 676-689.

Monterosso, J., Ehrman, R., Napier, K. L., O'Brien, C. P., & Childress, A. R. (2001). Three decision-making tasks in cocaine-dependent patients: Do they measure the same construct? *Addiction, 96*(12), 1825-1837.

Rotheram-Fuller, E., Shoptaw, S., Berman, S. M., & London, E. D. (2004). Impaired performance in a test of decision-making by opiate-dependent tobacco smokers. *Drug and Alcohol Dependence, 73*(1), 79-86.

Stout, J. C., Busemeyer, J. R., Lin, A., Grant, S. J., & Bonson, K. R. (2004). Cognitive modeling analysis of decision-making processes in cocaine abusers. *Psychonomic Bulletin & Review, 11*(4), 742-747.

Stout, J. C., Rock, S. L., Campbell, M. C., Busemeyer, J. R., & Finn, P. R. (2005). Psychological processes underlying risky decisions in drug abusers. *Psychology of Addictive Behaviors, 19*(2), 148-157.

Sutton, R. S., & Barto, A. G. (1998). *Reinforcement Learning: An Introduction*. Cambridge, MA: MIT Press.

پژوهشگاه علوم انسانی و مطالعات فرهنگی
پرتال جامع علوم انسانی