

# نظریه بازی‌ها به مثابه ابزاری برای فلسفه اخلاق؛ تبیینی تاریخی و تحلیلی انتقادی

تاریخ دریافت: ۱۳۹۵/۱/۲۷ تاریخ تأیید: ۱۳۹۵/۳/۲۵

سیده‌های عربی\* محمد‌های زاهدی و فاطمه\*\*  
محمدجواد رضائی\*\*\* مهدی موحدی بکنظر\*\*\*\*

## چکیده

فلسفه اخلاق، پژوهشی درباره درستی، نادرستی، روایی یا ناروایی اعمال یا تحقیق درباره استدلال‌های ویرای باورهای اخلاقی ماست. یکی از عرصه‌های مهم فراخوانی باورهای اخلاقی یا درک عمل درست یا نادرست، حوزه رفتارهای تعاملی و متقابل انسان‌هاست؛ از این‌رو منطقی به نظر می‌آید که نظریه بازی‌ها، به‌عنوان مطالعه فرایند و برابند تعامل‌های افراد، ظرفیت برخی کاوش‌های فلسفی و اخلاقی را داشته باشد. مقاله پیش رو در پی آن است که نخست اثرگذاری برخی ایده‌ها و مدل‌های نظریه بازی‌ها بر برخی ایده‌ها و نظریه‌ها در حوزه فلسفه اخلاق و عدالت پژوهی را بررسی کرده؛ سپس استواری اعتبار و کارآمدی چنین اثرهایی را تحلیل انتقادی کند. شیوه پژوهش در قسمت نخست مبتنی بر مطالعه‌های تاریخی از ۱۹۴۰م تاکنون است. ارزیابی انتقادی ارائه‌شده نیز به شکل تحلیل نظری است. یافته‌های پژوهش بیان‌گر آن است که برخی نوآوری‌های صورت‌گرفته در نظریه بازی‌ها بر برخی ایده‌ها و نظریه‌ها در حوزه فلسفه اخلاق و عدالت پژوهی اثرگذار بوده است. ارزیابی انتقادی چارچوب تحلیلی ایده‌ها و نظریه‌های مبتنی بر رهیافت نظریه بازی‌ها، نشان‌دهنده کاستی‌ها و دشواری‌هایی در راه بهره‌گیری از نظریه بازی‌ها به‌عنوان ابزاری برای فلسفه اخلاق است.

واژگان کلیدی: فلسفه اخلاق، فلسفه اقتصاد، عقلانیت، عدالت، نظریه بازی‌ها، نظریه بازی‌های تطوری.

طبقه‌بندی JEL: A12, C73, D63, D74

۹۱

فصلنامه علمی پژوهشی اقتصاد اسلامی / سال شانزدهم / شماره ۶۳ / پاییز ۱۳۹۵

Email: Sharabi@rihu.ac.ir.

\*. دانشیار دانشگاه قم.

Email: Zahedi@isu.ac.ir.

\*\* . دانشیار دانشگاه امام صادق (ع).

\*\*\* . دانشجوی دکتری علوم اقتصادی دانشگاه امام صادق (ع) (نویسنده مسئول).

Email: J.rezaei@isu.ac.ir.

Email: Movahedi@isu.ac.ir.

\*\*\*\* . دانشجوی دکتری علوم اقتصادی دانشگاه امام صادق (ع).

## مقدمه

تاریخ پژوهش‌های فلسفی سرشار از انواع اثرپذیری‌ها، اثرگذاری‌ها از/ بر حوزه‌های دیگر علوم و معارف است؛ انواع روابط میان «فلسفه اخلاق و علوم سیاسی»، «معرفت‌شناسی و علم اقتصاد» یا «فلسفه و عصب‌شناسی» بیانگر گستره قابل توجه چنین مطالعه‌هایی است. مقاله پیش رو قصد دقت کردن در یکی از این مناسبت‌ها را دارد. «فلسفه اخلاق» و «نظریه بازی‌ها» دو حوزه به ترتیب باسابقه و نوظهور در حوزه فعالیت‌های مطالعاتی و پژوهشی در علوم انسانی و علوم اجتماعی هستند که به نظر می‌رسد بنا به موضوع و مبادی‌شان عرصه‌هایی برای وجود ارتباط میان آنها قابل تصور است. مقاله پیش رو به آثار برخی پژوهشگران که در این عرصه گام برداشته‌اند توجه داشته و قصد بررسی دو مطلب ذیل را دارد:

الف) تأثیرپذیری برخی مطالعه‌ها در فلسفه اخلاق از مطالعه‌های نظریه بازی‌ها؛

ب) ارزیابی انتقادی از پروژه «نظریه بازی‌ها به مثابه ابزاری برای فیلسوف اخلاق».

گفتنی است کاوش درباره بحث‌های میان‌رشته‌ای نظریه بازی‌ها و فلسفه اخلاق به طور صریح یا ضمنی، بُعدهای مهمی از علم اقتصاد را در نظر دارد و از این رو است که بحث‌های مورد نظر در مقاله در ارتباط کامل با ادبیات فلسفه اقتصاد تلقی می‌شود (Hausman & McPherson, 2006)\*.

از این رو در ادامه اشاره‌ای به ماهیت فلسفه اخلاق و نظریه بازی‌ها داشته تا عرصه بحث درباره مناسبت‌های میان این دو حوزه آشکار شود. در بخش دوم مصداق‌هایی از اثرگذاری برخی ایده‌ها و مدل‌های نظریه بازی‌ها بر برخی ایده‌ها و نظریه‌ها در حوزه فلسفه اخلاق و عدالت ترسیم خواهد شد. در بخش نهایی مقاله نیز ارزیابی از پروژه مد نظر صورت خواهد گرفت.

## فلسفه اخلاق و نظریه بازی‌ها

دانش اخلاق یا مجموعه مطالعه‌ها و پژوهش‌ها درباره اخلاق را می‌توان در سه دسته کلی تقسیم کرد (مصباح یزدی، ۱۳۸۴، ص ۲۳-۲۴ / پالمر، ۱۳۸۸، ص ۱۵):

\*. این کتاب هاسمن و مکفرسون در چهار بخش تنظیم شده است که بخش چهارم آن با عنوان ریاضیات اخلاق (Moral Mathematics) دربرگیرنده فصل‌های سیزده و چهاردهم به ترتیب با عنوان‌های «نظریه انتخاب اجتماعی» و «نظریه بازی‌ها» است.

۱. برخی مطالعه‌ها به پژوهش‌های توصیفی از انواع اخلاقیات در جوامع یا افراد گوناگون می‌پردازد. در این دسته، هدف روایت بحث‌های اخلاقی بدون هر گونه ارزش‌دوری است.

۲. برخی پژوهش‌ها ذیل اخلاق هنجاری یا همان نظریه‌های اخلاقی دسته‌بندی می‌شوند. در این گونه از مطالعه‌ها هدف آن است که ملاک‌های اتصاف اعمال، نتیجه‌ها، ویژگی‌های شخصی یا محرک‌های انتخاب به خوب و بد یا درست و نادرست بحث شود.

۳. فرااخلاق به تحلیل فلسفی درباره معنا و ویژگی زبان اخلاق می‌پردازد. در این دسته از پژوهش‌ها بُعدهای معناشناختی، وجودشناختی و معرفت‌شناختی مفهوم‌های اخلاقی - مانند خوب و بد، روا و ناروا - بررسی می‌شود.

به بیان ساده، فلسفه اخلاق شامل مجموعه مطالعه‌های این دو دسته واپسین است؛ مجموعه مطالعه‌هایی درباره روایی یا ناروایی اعمال، خوبی یا بدی نتیجه‌ها، فضیلت‌مندی یا رذیلت‌مندی ویژگی‌های شخصی یا خیر و شر بودن محرک‌ها و اراده عمل. واضح است که اخلاق توصیفی به کلی از مجموعه مطالعه‌های دو دسته واپسین متفاوت است. گفتنی است همان‌طور که از تعریف فلسفه اخلاق بر می‌آید، عدالت به عنوان یکی از مهم‌ترین جستارهای آن قابل طرح است؛ مقوله‌های ارزشی یا اخلاقی مورد بحث در فلسفه اخلاق شامل مفهوم‌ها یا فضیلت‌های مهمی مانند آزادی، عدالت، حق، اراده، فایده (Utility)، تکلیف، نیت و ... است؛ از این رو هر بررسی و پژوهشی درباره یکی از مقوله‌های پیش‌گفته، ذیل پژوهش‌های فلسفه اخلاق می‌گنجد؛ مانند دانش اخلاق که درباره اعمال، انتخاب‌ها، ویژگی‌ها یا محرک‌ها پژوهش می‌کند، حوزه شاخصی در مجموعه علوم اجتماعی وجود دارد که موضوع خود را مطالعه درباره برخی انتخاب‌های انسان کرده است. علم اقتصاد به بررسی «انتخاب (Choice)»، «رفتارهای ناظر به مبادله (Exchange) انسان‌ها» یا «تخصیص منابع محدود به خواسته‌های نامحدود» می‌پردازد. به نظر می‌آید با اتصاف این مجموعه مطالعاتی به علم (Science) در ظاهر جایگاهی برای بحث‌های هنجاری در آن وجود ندارد. از این دید و در مقام تشبیه، علم اقتصاد را می‌توان فعلاً و از روی تسامح با اخلاق توصیفی متناظر دانست؛ اینکه انسان‌ها در زندگی روزمره خود چگونه تصمیم گرفته یا دست به انتخاب‌های اقتصادی می‌زنند و نه اینکه باید چگونه دست به انتخاب بزنند.

هنگامی که بحث انتخاب پیش می‌آید، می‌توان بسترهای گوناگونی برای بروز این انتخاب متصور بود. گاهی اوقات مسئله علم اقتصاد توجه به تصمیم یا انتخاب انسان در

وضعیتی مانند وضعیت رقابتی بازار آزاد است؛ در این وضعیت مؤلفه‌های مهم در انتخاب فرد، مجموعه هزینه‌ها و فایده‌های به دست آمده از گزینه‌های پیش رو است. در این چارچوب تحلیلی، اثر تک‌تک افراد در فرایند مبادله آزاد در بازار آن‌قدر ناچیز است که در عمل هیچ تک‌فردی در نتیجه‌های تخصیصی بازار اثرگذار نیست.

یکی از انتقادهای وارد بر چنان شیوه تحلیلی در علم اقتصاد، در نظر نگرفتن فعالیت‌هایی است که اتفاقاً انتخاب افراد بر نتیجه‌های تعامل با یکدیگر اثرگذار بوده و به بیان دقیق‌تر نتیجه‌های انتخاب یک فرد، تابعی از انتخاب فرد یا افراد مقابل است؛ نمونه ساده آن، موقعیتی است که با عنوان معمای زندانی (The prisoner's dilemma) شناخته می‌شود. در این وضعیت، نتیجه نهایی برای یکی از بازیکنان، تابعی از انتخاب زندانی دوم است. توجه به چنین موقعیت‌های تعاملی و تفاوت آن با تحلیل‌های مبتنی بر بهینه‌سازی‌های متداول در علم اقتصاد بود که بستر را برای ظهور نظریه بازی‌ها فراهم آورد.

تا اینجای بحث به دریافتی اجمالی از جستارمیه‌های دانش اخلاق، فلسفه اخلاق، علم اقتصاد و نظریه بازی‌ها دست یافتیم. اکنون نوبت به درک ارتباط میان فلسفه اخلاق و نظریه بازی‌ها می‌رسد. به نظر می‌رسد مهم‌ترین حلقه واسط میان فلسفه اخلاق و نظریه بازی‌ها مقوله عقلانیت (Rationality) است. در ادامه به جایگاه این مفهوم در فلسفه اخلاق و نیز نظریه بازی‌ها می‌پردازیم.

عقلانیت یکی از مفهوم‌ها یا واژه‌های مهم در بسیاری از نظریه‌های اخلاقی و پژوهش‌های فلسفه اخلاق است تا حدی که به بیان برخی در اخلاق هنجاری می‌کوشیم برای اینکه بتوانیم با ابزار عقلانی به مجموعه‌ای از معیارهای قابل قبول دست یابیم تا ما را در تعیین اینکه چرا فعل خاصی «صواب» یا شخص خاصی «خوب» خوانده می‌شود قادر سازد (پالمر، ۱۳۸۸، ص ۱۹). از این رو است که از دید برخی فیلسوفان اخلاق، درک اصول اخلاق به واسطه عقل یا مبتنی بر منطق اکتشاف عقلایی صورت می‌گیرد. شاید بتوان از اختلاف میان هیوم و کانت در این موضوع به عنوان شاخص‌ترین نمونه یاد کرد. هیوم اخلاق را امری برانگیخته از احساس‌ها می‌داند و در برابر، حوزه نفوذ و کارکرد دیگری برای عقل قائل است. کانت اخلاقی‌بودن عملی را انطباق آن با اصولی عقلایی و قابل استدلال می‌داند. به بیان رالز که خود را از قراردادگرایان و کانت متأثر می‌داند، نه تنها سنت

قراردادگرایی بلکه فلسفه اخلاق بخشی از نظریه انتخاب عقلایی است ( Rawls, 2005, pp.16, 47, 172).

همان‌طور که اشاره شد، علم اقتصاد و نیز نظریه بازی‌ها به مطالعه انتخاب انسان می‌پردازد. ابزار اقتصاددانان و نیز متخصصان نظریه بازی‌ها در این راه، مدل‌سازی رفتار انسان است. آنان مدعی هستند یکی از ویژگی‌های هر نوع مدل‌سازی، ساده‌سازی واقعیت مورد مطالعه است؛ از این‌رو انسان مدل‌سازی شده در علم اقتصاد، نوعی انسان آرمانی با ویژگی عقلایی بودن است. در علم اقتصاد و نیز نظریه بازی‌ها، این انسان اقتصادی یا انسان عقلایی است که بار تبیین‌های اقتصاددانان و نیز پیش‌بینی‌هایشان را به دوش می‌کشد.

البته باید توجه داشت تلقی‌ای که از مفهوم عقلانیت در مطالعه‌های اقتصادی وجود داشته و دارد، تطور مهمی یافته است. از تلقی اولیه از مفهوم عقلانیت در علم اقتصاد بر می‌آمد که این مفهوم دربرگیرنده برخورداری «معیاری برای انتخاب» بوده که به پیروی از آن به عنوان صفتی برای عمل یا کنش استفاده می‌شود؛ ترکیب‌های انتخاب عقلایی، گزینش عقلایی یا تصمیم عقلایی همگی بر این اساس معنادار به نظر می‌آیند. برای معناداری این ترکیب‌ها، آنچه مهم است وجود ملاک و معیاری برای اتصاف یک انتخاب، گزینش یا تصمیم به صفت عقلایی است. یگانه معیار عقلانیت در این تلقی نیز همان پیگیری منافع شخصی یا پیگیری بیشترین فایده (مطلوبیت) برای فرد یا افراد است.

با توجه به گرایش اقتصاددانان به علمی دانسته‌شدن رشته آموزشی خویش و نیز وجود دشواری‌هایی در تلقی هنجاری از عقلانیت، به مرور زمان برخی اقتصاددانان از اصرار ورزیدن بر تعریف‌های محتوایی از عقلانیت دست کشیده و تلقی دیگری از عقلانیت را که قابلیت فرموله‌شدن داشته باشد را به کار بردند. تلقی صوری از عقلانیت دیگر در مقام ارائه راهنمای عمل در تمام موقعیت انتخاب نبوده و فقط از سازگاری انتخاب‌های فردی به عنوان معیار عقلایی بودن انتخاب‌های وی یاد می‌کند؛ به طور مثال، در این حالت مهم نیست که فرد از میان سیب، موز و پرتقال کدام را انتخاب می‌کند؛ آنچه مهم است آن است که فرد در انتخاب‌های خویش از میان سیب، موز و پرتقال همواره پایبند به شرط‌های کامل بودن (Completeness) و انتقال‌پذیری (Transitivity) باشد؛ یعنی اگر در یک

موقعیت انتخابی از میان این سه بدیل، سیب را به دو میوه دیگر ترجیح داد، همواره این ترجیح در رفتار وی مشاهده شود و نیز اگر سیب را به موز و موز را به پرتقال ترجیح داد، همواره سیب را به پرتقال نیز ترجیح دهد. واضح است که مبتنی بر این تلقی، در اساس نظریه انتخاب عقلایی به تمسک به هیچ نوعی از عامل‌های روان‌شناختی نیز نیازمند نیست؛ چرا که بر یک تلقی صوری از رفتار انتخابی سازگار مبتنی است (Lehtinen & Kuorikoski, 2007, p.119).

### نظریه بازی‌ها به مثابه ابزاری برای فیلسوف اخلاق

از دید برخی فیلسوفان یا مکتب‌های اخلاقی، عدالت یا درک اصول اخلاقی از رهگذر نظریه انتخاب عقلایی می‌گذرد. از طرف دیگر، نظریه بازی‌ها به بررسی انواع تعامل‌هایی پرداخته که البته در بسیاری از موارد فرض عقلانیت نیز در بازیکنان مورد مطالعه لحاظ می‌شود. حال که عدالت ذیل نظریه انتخاب عقلایی بوده و نظریه بازی‌ها هم به نوعی به مطالعه تعامل‌های بازیکنان عقلایی می‌پردازد، دور از انتظار نخواهد بود که ظهور، تغییر یا توسعه مفهوم‌ها و شیوه‌های تحلیل در نظریه بازی‌ها بر برخی نظریه‌پردازان در حوزه دانش اخلاق و به طور خاص فلسفه اخلاق اثرگذار بوده باشد.\* در این بخش با جزئیات بیشتری به تبیین این فرضیه مبنی بر رابطه میان برخی نوآوری‌های صورت‌گرفته در نظریه بازی‌ها و برخی نظریه‌ها درباره اخلاق و عدالت، به‌عنوان فرضیه نخست مقاله می‌پردازیم.

۱. بی‌شک گام نخست در بررسی این فرضیه با لحاظ کردن طرح اولیه ایده نظریه بازی‌ها با چاپ نخست نظریه بازی‌ها و رفتار اقتصادی در سال ۱۹۴۴م آغاز می‌شود. جان فون نویمان

\*. با گذر زمان، غلبه رویکردهایی مانند نظریه بازی‌ها در پژوهش‌های اخلاقی، اعم از اخلاق توصیفی و فلسفه اخلاق در حال افزایش است. کوشش برای وصف رفتارهای اخلاقی انسان‌ها بر اساس مدل‌های نظریه بازی‌ها امری است که نمی‌تواند جدا از بحث‌های فلسفه اخلاق دانسته شود، همان‌طور که در مقاله پیش رو نیز پافشاری بیشتر بر ارتباط میان نظریه بازی‌ها و بحث‌های اخلاق هنجاری است؛ اما اگر مدل‌های نظریه بازی‌ها به مثابه ابزارهایی برای وصف اخلاقیات انسانی نیز در نظر گرفته شود، مطالعه‌های بسیاری در این باره قابل رصد کردن است (Hoffman, Yoeli & Navarrete, 2016).

(John von Neumann) و اسکار مورگنشترن (Oskar Morgenstern) با کوشش برای طرح تحلیل رفتارهای تعاملی از بازی‌های بسیار ساده با جمع صفر (Zero-sum Game) آغاز کردند. آنها روش حداکثرسازی حداقل‌ها (Maximin) را به‌عنوان شیوه دستیابی به موقعیت‌های تعادلی معرفی کردند؛ روشی که تا زمان طرح نقطه تعادلی نش در ۱۹۵۱م، یگانه ابزار تحلیلی در دست پژوهشگران حوزه نظریه بازی‌ها بود.\*

طرح این شیوه حل موقعیت‌های تعاملی و دست‌یافتن به موقعیت تعادلی، طبعاً می‌تواند مورد توجه برخی از فیلسوفان قرار گیرد. اینجاست که می‌توان از رالز به‌عنوان نخستین فیلسوف متأثر از برخی مؤلفه‌های نظریه بازی‌ها یاد کرد. وی در نظریه عدالت در پی توجیه‌کردن اصول عدالت خویش می‌رود. آنچه شاهد مثال ماست، اصل تفاوت (The Difference Principle) وی و اثرپذیری وی از قاعده حداکثرسازی حداقل‌هاست. با طرح ایده رالز در قالب اصل تفاوت و پشتیبانی آن از سوی قاعده حداکثرسازی حداقل‌ها در موقعیت نخستین (Rawls, 2005, p.155) گام مهمی در شکل‌گیری نظریه عدالت وی برداشته شد. رالز تمسک به این قاعده را در موقعیت فرضی خویش و در پس پرده جهل، انتخابی عقلایی برای هر گزینش‌گری می‌داند.

توضیح آنکه روش حداکثرسازی حداقل‌ها از نظر اقتصاددانان و متخصصان نظریه بازی‌ها فقط ابزاری خاص و محدود برای تحلیل نوع خاصی از بازی‌هاست. با خارج شدن بازی از نوع با مجموع صفر، تحلیل‌ها معطوف به یافتن تعادل نش شده و دیگر روش حداکثرسازی حداقل‌ها کاربردی ندارد. به بیان دیگر قاعده حداکثرسازی حداقل‌ها حالت خاصی از یافتن استراتژی‌های تعادلی نش است.\*\*

وجود همین کاستی در قاعده حداکثرسازی حداقل‌هاست که برخی اقتصاددانان معاصر با رالز بر سر این شیوه تحلیل و در نتیجه عقلایی دانسته شدن انتخاب اصل تفاوت در پس پرده

---

\*. به بیان امروزی، روش حداکثرسازی حداقل‌ها برای یافتن یک تعادل نش در یک بازی با مجموع صفر استفاده می‌شود؛ با یافتن maximin و minimax و مشاهده یکسان بودن این مقادیر می‌توان آن نقطه را برخوردار از تعادل نش دانست (Dixit & Skeath, 2004, p.638).

\*\* به بیان کتاب‌های آموزشی نظریه بازی‌ها، بازی‌هایی که در واقعیت با آنها روبرو هستیم، کم‌وبیش هیچ‌گاه بازی‌های مجموع صفر نبوده و از این رو قضیه حداکثرسازی حداقل‌ها کاربست‌پذیری محدودی داشته و توصیه عمومی به عمل بر اساس آن غیرعقلانی است (Binmore, 2007b, p.233/ Rasmusen, 2005, p.145).

جهل مخالفت کردند. بنام‌ترین آنان، جان هارزانی (John Harsanyi) بود که این مخالفت به جدال مفصلی بین وی و رالز انجامید. از نظر وی، وجود برخی متناقض‌نماها و تغایر اصل تفاوت با تصمیم‌گیری عقلایی، دلیل موجهی برای نپذیرفتن اصل تفاوت در موقعیت نخستین است؛ چرا که انتظار بر آن است که آنها - افراد حاضر در موقعیت نخستین - بر اساس شیوه‌ای عقلایی (Rational Manner) عمل خواهند کرد؛ یعنی مبتنی بر قاعده‌های تصمیم‌گیری که به تصمیم‌های معقول در وضعیت نادانی و نااطمینانی بینجامد (Harsanyi, 1975, p.596). مبتنی بر تلقی هارزانی از انتخاب عقلایی، اعطای اولویت مطلق به نفع بدترین فرد، [آن هم] به هر شکلی که باشد قابل پذیرش نیست (Ibid, p.596).

خلاصه آنکه اثرپذیری ضمنی یا صریح رالز از ایده حداکثرسازی حداقل‌ها بود که عرصه را برای برخی اقتصاددانان چون هارزانی فراهم کرد که به اعتراض علیه طرح عدالت رالز برخیزند.\*

\*. ممکن است این پرسش پیش آید که مگر نه آنکه رالز خود را از کانت متأثر می‌داند و به نوعی در برابر تحلیل‌های فایده‌گرایانه از عدالت، قد علم کرده است؟ پس چگونه از یک ایده در نظریه بازی‌ها متأثر شده است؟ پاسخ به این پرسش نیازمند دقت و ورزی بیشتری در شیوه تحلیل و توجیه رالز است. به نظر می‌رسد آنجا که رالز راجع به اصل تفاوت و شیوه‌گزینش اصول عدالت در پس پرده نادانی استدلال می‌کند، اصلاً شباهتی با تکلیف‌گرایی کانتی وجود نداشته و وی به طور کامل در چارچوب سنت فایده‌گرایی به طرح بحث می‌پردازد. اساساً همین اعتنا و اتکای وی به فایده‌گرایی است که باعث می‌شود هارزانی به عنوان یک فایده‌گرا، وارد مجادله با وی شود. جالب آنکه رالز در جایی نیز به صراحت مبنای گزینش افراد در پس پرده نادانی را چیزی شبیه قاعده «هرچه بیشتر بهتر» (The more is better) بیان می‌کند. به بیان رالز، راجع به افراد حاضر در پس پرده نادانی «... بر این فرض هستند که خیرهای اساسی اجتماعی بیشتر را به کمتر ترجیح می‌دهند» (Rawls, 2005, p.142). به بیان دیگر، مطالبه سهم بیشتری از خیر برای طرف‌ها عقلایی است چرا که سرانجام و پس از افتادن پرده نادانی، حتی هم اگر یکی از طرف‌ها سهمی که بیش از تمایل خودش بوده را به دست آورده باشد به استفاده از مجبور آن نیست؛ در نتیجه می‌توان اعلام کرد که دست‌کم در بخشی از توجیه‌های ارائه‌شده در نظریه عدالت، مبنای انسان‌شناسی رالز تفاوتی با مبنای انسان‌شناختی نظریه‌های فایده‌گرا یا نتیجه‌گرا ندارد. همچنین کافی است با دقت بیشتری به ایده رالز راجع به گزینش اصول عدالت در پس پرده نادانی توجه کنیم: «اینها - یعنی اصول عدالت - اصولی هستند که افراد عاقل به منظور پی‌گیری منافع خود آنها را در این موقعیت - نخستین - برابری به‌منظور بنا نهادن مفهوم‌های بنیادین مجمع خود می‌پذیرند» (Ibid, pp.118-119).



۲. یکی دیگر از نوآوری‌ها در مطالعه‌های نظریه بازی‌ها، در حوزه موقعیت‌های چانه‌زنی (Bargaining) بود. مدل‌های چانه‌زنی در پی آن بودند که بر اساس برخی آگزیوم‌ها، راه‌حلی برای موقعیت‌های چانه‌زنی فراهم آورند؛ به طور مثال، موقعیت‌های تسهیم یک منبع بادآورده میان دو نفر. نخستین راه‌حل پیشنهادی در این حوزه نیز به وسیله جان نش (John Forbes Nash, 1950) ارائه شد. این حوزه از مطالعه‌های متخصصان نظریه بازی‌ها به مرور زمان راه‌حل‌های پیشنهادی دیگری را نیز به خود دید.

دیوید گوتیر (David Gauthier) \* با «اخلاقیات بر اساس توافق» (۱۹۸۶م) به عنوان طراح ایده‌ای درباره اخلاق و عدالت در چارچوب تحلیل‌های مبتنی بر بازی‌های ساده در نظریه بازی‌ها و البته در چارچوب رهیافت‌های قراردادگرایانه شناخته می‌شود. گوتیر با تعریف موقعیت اولیه برای چانه‌زنی، در پی ارائه راه‌حلی بر اساس مدل‌های متعارف چانه‌زنی بود؛ شیوه‌ای که به‌طوری عقلایی نه‌تنها ارائه‌کننده ملاکی برای دستیابی به توافقی اخلاقی باشد بلکه در مسائلی مانند عدالت توزیعی نیز استفاده شود. \*\*

یکی از اجزای طرح گوتیر، پیشنهاد راه‌حلی برای موقعیت‌های چانه‌زنی در نظریه بازی‌ها بود؛ راه‌حلی که از نظر خودش پاسخی عقلایی و منحصربه‌فرد برای چنین موقعیت‌هایی است (Gauthier, 1986, p.139). جالب آنکه راه‌حل پیشنهادی گوتیر، گونه‌ای از راه‌حل چانه‌زنی کالای - / سموردینسکی (۱۹۷۵م) بوده که مدت‌ها پیش پیشنهاد شده بود که البته خود گوتیر هم به این اثرپذیری اشاره کرده است (Ibid. p.130). گفتنی است که وی به‌واسطه برخی انتقادات، بعدها راه‌حل پیشنهادی خویش را به کناری نهاده و

\*. فیلسوف کانادایی - امریکایی که به‌عنوان فیلسوف اخلاق نئوهابزی شناخته شده است. مهم‌ترین کتاب وی **اخلاقیات بر اساس توافق** (۱۹۸۶م) است که توجه بسیاری در مجامع فلسفی، فلسفه اخلاق و فلسفه سیاسی را به خود جلب کرد.

\*\* گفتنی است که از دید بسیاری از متخصصان میان‌رشته‌ای در حوزه فلسفه و اقتصاد، رفتار اخلاقی معادل بروز هم‌آهنگی و همکاری میان بازیکنان است (به‌طور مثال در Curry, 2016). از این دید اخلاق اساساً در موقعیت‌های تعاملی معنادار بوده و قرار بر آن است که به نحوی از انحاء به حل تعارض منافع میان بازیکنان پرداخته شود.

به راه حل چانه‌زنی نش؛ یعنی راه‌حلی که بیش از چهل سال پیش در ادبیات نظریه بازی‌ها پیشنهاد شده بود، مراجعت کرد (Gauthier, 1993, p.180).\*

۳. تأییدکننده دیگر به طرح قضیه عامه (Folk Theorem) در بازی‌های تکراری و اثرپذیری بخشی از عناصر استدلالی کنت بینمور (Kenneth Binmore)\*\* در نظریه عدالتش بر می‌گردد. در دهه شصت میلادی (بازه ۱۹۵۰م - ۱۹۶۰م) قضیه عامه در بازی‌های تکراری میان متخصصان نظریه بازی‌ها معروف و آشنا بوده (Gibbons, 1992, p.89) و ابداع یا اثبات آن متعلق به فرد خاصی نبود. نخستین تبیین‌های بینمور از نظریه خویش نیز با فاصله‌ای قریب به سی سال از طرح قضیه عامه به این قضیه مهم در بازی‌های تکراری بر می‌گردد. بینمور در کتاب دو جلدی خویش با عنوان اصلی نظریه بازی‌ها و قرارداد اجتماعی (۱۹۹۴م و ۱۹۹۸م) و نیز کتاب متأخر خویش با عنوان عدالت طبیعی (۲۰۰۵م) به طرح ایده خویش درباره عدالت پرداخته است.

مسئله اصلی بینمور مانند هابز و نیز رالز، شیوه دستیابی به قراردادی اجتماعی عادلانه است. وی سه شرط برای چنین قراردادی برمی‌شمارد: الف) پایداری، ب) کارایی، ج) منصفانه‌بودن. محل تمرکز ما همان شرط نخست وی است. بنا به تحلیل وی، مناسب‌ترین ابزار برای تحلیل موقعیت پایدار یا همان تعادلی، تحلیل از دید بازی‌های تکراری است. قضیه عامه نیز که برای بازی‌های تکراری صادق بوده و تعیین‌کننده استراتژی‌های تعادلی بازیکنان در این وضعیت می‌باشد، تعیین‌کننده قراردادهای اجتماعی پایدار برای وی است.

۴. در گام نهایی نیز باید به طرح نظریه بازی‌های تطوری (Evolutionary Game Theory) اشاره کنیم. خاستگاه نظریه بازی‌های تطوری را می‌توان بهره‌گیری از نظریه بازی‌ها در فضای علوم

\*. در متن‌های فلسفی و فلسفه اخلاق به زبان فارسی اشاره‌های اندکی به سیر استدلال‌های گوتیر و پروژه وی شده است. منبع زیر مقاله حاصل از جستار نویسندگان بود که به نظر درآمد خوبی برای ورود به زمینه‌های فکری گوتیر است (پیک حرفه، ۱۳۹۰، ص ۶۹ - ۸۸).

\*\* اقتصاددان و متخصص نظریه بازی‌های بریتانیایی است. وی یکی از نوآوران در حوزه نظریه چانه‌زنی در نظریه بازی‌هاست. به‌واسطه مجموعه کارهایی که در حوزه‌های میان‌رشته‌ای اقتصاد، علوم سیاسی و اخلاق انجام شده است، وی را از فیلسوفان تحلیلی می‌شمارند.

زیستی و مطالعه تحول‌های جمعیتی در محیط‌های با ویژگی تقابلی و البته غیرارادی دانست. شکل‌گیری چنین رویکردی متأثر از ارائه نظریه بازی‌ها به وسیله فون نویمان و مورگن‌شترن و انجام برخی مطالعه‌ها به وسیله زیست‌شناسان و ریاضی‌دانان بوده است. شکل‌گیری نظریه بازی‌های تطوری بیش از همه، مدیون جان مینارد اسمیت است. مهم‌ترین مفهوم در نظریه بازی‌های تطوری یعنی استراتژی پایدار تطوری، نخستین بار به وسیله مینارد اسمیت در ۱۹۷۲م صورت‌بندی شده؛ سپس به شکل رسمی‌تر به همراه یکی از همکارانش تبیین شد (Smith & Price, 1973). به‌طور ساده نیز منظور از استراتژی پایدار تطوری، حالتی است که ترکیب جمعیتی بازیکنان با انتخاب آن استراتژی به واسطه پیدایش جهش‌های احتمالی تغییر نکند.

این نظریه که به نوعی طرحی نو در مجموعه مطالعه‌های نظریه بازی‌ها بود با فاصله‌ای نزدیک به پانزده سال، نظر رابرت ساگدن (Roert Sugden, 1989)\* را به خود جلب کرد و به‌طور آشکاری در آثار برایان اسکرمز (Brian Skyrms, 1996)\*\* به اوج بهره‌گیری در موضوع عدالت یا انصاف دست یافت. این افراد در موضوع‌هایی چون مالکیت یا تسهیم منابع بادآورده، بر اساس تحلیل‌های نظریه بازی‌های تطوری به منصفانه دانسته‌شدن تقسیم

---

\*. پژوهشگر بریتانیایی که در دو حوزه اقتصاد رفتاری و فلسفه سیاسی آثار قابل توجهی دارد. هر چند وی به‌عنوان فیلسوف اخلاق شناخته نمی‌شود و بیشتر به‌عنوان اقتصاددان یا فیلسوف اقتصادی مطرح است اما انتشار برخی آثارش در مجله‌های فلسفه اخلاق و نیز توجه مجامع و مجله‌های اصیل فلسفه اخلاق به مهم‌ترین کتاب وی در حوزه مورد بحث؛ یعنی اقتصاد حق، همکاری و رفاه (۲۰۰۵م [۱۹۸۶])، باعث شد ایده وی در حوزه فلسفه اخلاق را برخوردار از صلاحیت لازم برای بررسی بدانیم. منابع زیر به‌عنوان دو نمونه از توجه به آثار وی، قابل اشاره هستند (-Verbeek, 2010, pp.541/ Pettit, 1990, pp.725-755, 559).

\*\* وی استاد منطق، فلسفه و اقتصاد دانشگاه کالیفرنیا در اروین و نیز استاد فلسفه دانشگاه استنفورد است. مهم‌ترین حوزه‌های پژوهشی وی نظریه تصمیم‌گیری، مبانی احتمال‌ها و نظریه بازی‌هاست. کارهای پژوهشی متأخر وی از جمله دو کتاب **تطور قرارداد اجتماعی** (۱۹۹۶م و نیز تجدید چاپ در ۲۰۱۴م) و **شکار گوزن و تطور ساختار اجتماعی** (۲۰۰۴م)، درباره ماهیت قرارداد اجتماعی با استفاده از نظریه بازی‌های تطوری است. به بیان هاسمن و مکفرسون، اسکرمز کاربردهایی شگرف از نظریه بازی‌های تطوری در حوزه هنجارهای اجتماعی را اظهار داشته است (هاسمن و مکفرسون، ۲۰۰۶م. ص ۲۵۵).

۵۰-۵۰ در نظر افراد و نیز پایدارماندن این شیوه عمل نظر داشته و اساس تحلیل خویش از هنجارها، اصول یا میثاق‌های اجتماعی عادلانه را وجود چنین مشخصه‌ای قرار دادند.

ساگدن با مهم دانستن امکان بروز هم‌آهنگی میان افراد آن هم بدون وجود هرگونه قوه قاهره‌ای بر استعاره نظم خودانگیخته دست می‌گذارد. ساگدن با استفاده از مفهوم تعادل به تحلیل سه میثاق مهم هم‌آهنگی، مالکیت و معامله به مثل می‌پردازد. از نظر وی میثاق‌های هم‌آهنگی قابل مشاهده در زندگی روزمره که بدون وجود هرگونه قوه قاهره‌ای بروز می‌کند، ابزاری برای هم‌آهنگی و بروز موقعیت تعادلی است. وی برای تبیین میثاق مالکیت از ابزار نظریه بازی‌های تطوری استفاده می‌کند و پیدایش میثاق‌هایی مانند «اگر مالک چیزی هستی تا سرحد مرگ برای حفظش بجنگ (استراتژی باز) و اگر مالک چیزی نیستی درگیر نشو (استراتژی کبوتر)» را استراتژی‌های پایدار تطوری دانسته و البته با ابزارهای مربوطه نیز در اثبات پایداری آن می‌کوشد.

اسکرمز با استفاده از بازی تقسیم عایدی اولیه که از چندین تعادل برخوردار هستند، موقعیت‌های دارای وضعیت تعادل پایدار تطوری را وضعیت عادلانه تلقی می‌کند. وی در طرح ایده خویش پافشاری بسیاری بر انتخاب بازی متناسب با موضوع قرارداد اجتماعی داشته و از این‌رو بازی شکار گوزن (Stag Hunt)\* را برای این منظور بر می‌گزیند. وی با شروع از وضع طبیعی هابزی در تعادل غالب ریسکی، فرایند تغییر موقعیت جامعه به وضعیت غالب پیامدی را رصد می‌کند.\*\*

سرانجام آنکه با توجه به شواهد پیش‌گفته و امکان رصد چنین رابطه‌ای می‌توان فرضیه وجود رابطه‌ای از طرف برخی پیشرفت‌ها در مجموعه مطالعه‌های نظریه بازی‌ها به برخی ایده‌های معاصر درباره عدالت را تأیید کرد. دلالت این فرضیه آن است که بیان می‌دارد

\*. این بازی ساده مانند معمای زندانی از بازی‌های ابزاری (Toy Games) تلقی می‌شود که برای تحلیل ماهیت قرارداد اجتماعی به وسیله اسکرمز پافشاری شده است. گفتنی است این بازی برگرفته از ایده ژان ژاک روسو راجع به قرارداد اجتماعی است.

\*\* منظور از تعادل غالب ریسکی وضعیتی است که همه افراد جامعه در پی شکار خرگوش - کار فردی که به طور قطع عایدی در برخواهد داشت - می‌روند. منظور از تعادل غالب پیامدی هم آن است که افراد دست به همکاری زده و به شکار گوزن که البته پیامد بالاتری برای آنان فراهم می‌آورد، دست می‌زنند.

برخی نظریه‌های معاصر طرح شده درباره عدالت و اخلاق با تأخیری زمانی متأثر از نوآوری‌های صورت گرفته در نظریه بازی‌ها هستند (رسم توضیحی یک).

• طرح بازی‌های با جمع صفر و پیشنهاد راه‌حل **maximin** به وسیله توسط فون نویمان و مورگنشرن (۱۹۴۴[۱۹۵۳]م)

• استفاده از قاعده **maximin** به عنوان قاعده انتخاب عقلانی برای گزینش اصول عدالت توسط رالز (۱۹۷۱[۲۰۰۵]م)

۱

• طرح مدل‌های چانه‌زنی با پیشگامی نش (۱۹۵۰م) و نوآوری‌هایی توسط کالای و اسموردینسکی (۱۹۷۵م)  
• طرح ایده اخلاق مبتنی بر توافق با استفاده از مدل‌های چانه‌زنی توسط گوتیر (۱۹۸۶م) و نیز تأثیرپذیری از راه‌حل چانه‌زنی در بازی‌های ایده خویش در باب اخلاق و عدالت (۱۹۹۳م)

۲

• قضیه عامه در بازی‌های تکراری (شروع طرح و بهره‌گیری از آن مشخص نیست، به قبل از ۱۹۶۰م برمی‌گردد)  
• طرح ایده عدالت طبیعی با استفاده از قضیه عامه در بازی‌های تکراری توسط یینمور (۱۹۹۴م)

۳

• طرح نظریه بازی‌های تطوری و ابداع مفهوم **ESS** با پیشگامی مینارد اسمیت و پرایس (۱۹۷۳م)  
• طرح ایده ساگدن (۱۹۸۹م)، اسکرمز (۱۹۹۴م) و یینمور (۱۹۹۴م) راجع به هنجارهای انصاف، عدالت، قرارداد اجتماعی و میناق‌ها بر پایه نظریه بازی‌های تطوری

۴

رسم توضیحی یک: اثرپذیری طرح برخی نظریه‌های مربوط به اخلاق، عدالت یا قرارداد اجتماعی از ظهور مفهوم‌ها و راه‌حل‌های جدید در مجموعه مطالعه‌های نظریه بازی‌ها

## نجات فلسفه اخلاق از نظریه بازی‌ها

حال نوبت به ارزیابی انتقادی از اثرگذاری چارچوب تحلیلی و مدل‌های نظریه بازی‌ها بر مطالعه‌های فلسفه اخلاق می‌رسد. دغدغه اصلی در این بخش فراهم آوردن پاسخ‌هایی برای پرسش‌های زیر است: «آیا نظریه بازی‌ها ابزار مناسبی برای فیلسوفان اخلاق است؟» یا «آیا

می‌توان به نظریه بازی‌ها به‌عنوان ابزاری خالی از محتوا نگریست که قابلیت به‌کارگیری به‌وسیله نظریه‌های اخلاقی گوناگون را دارد؟»

پیش از پی‌گیری بحث‌ها، لازم است تفکیکی میان دو نوع چارچوب تحلیلی در نظریه بازی‌ها قائل شده و پاسخ به پرسش‌های پیش‌گفته را در هر دوی این حوزه‌ها، به‌طور مجزا مد نظر قرار دهیم. سه دسته نوآوری نخست در حوزه نظریه بازی‌ها که بحث شد - قاعده حداکثرسازی حداقل‌ها، مدل‌های چانه‌زنی و قضیه عامه در بازی‌های تکراری را می‌توان در حوزه مطالعه‌های نظریه بازی‌های کلاسیک تلقی کرده و از نظریه بازی‌های تطوری - حوزه چهارم اشاره‌شده در بخش پیش - تفکیک کرد. علت لزوم این تمایز به تفاوت مفروض‌ها و شیوه‌های تحلیلی این دو دسته بر می‌گردد.

چند تفاوت عمده بین این دو رویکرد دیده می‌شود: نخست اینکه ادعا می‌شود نظریه بازی‌های تطوری بر خلاف نظریه بازی‌های کلاسیک به تحلیل کارهای توده‌ای و جمعیتی می‌پردازد. تفاوت دیگر اینکه در نظریه بازی‌های کلاسیک، مسئله انتخاب مبتنی بر بیشینه‌سازی فایده فردی صورت می‌گیرد؛ در حالی که در نظریه بازی‌های تطوری، انتخاب هوشمندانه‌ای دخیل نبوده بلکه انسان‌ها به مثابه موجودهایی خواهند بود که در مقام تحلیل، گویی‌گزینه‌ش طبیعی (Natural Selection) - داروینی - مبتنی بر تناسب (Fitness) - توانایی زیستی - بالاتر بر رفتار آنها حاکم است. تفاوت اجمالی این دو رهیافت به طور اجمالی در قالب جدول یک اشاره شده است؛ در نتیجه در دو قسمت بعد به تفکیک و ترتیب به دو حوزه پیش‌گفته خواهیم پرداخت.

جدول یک: تفاوت‌های نظریه بازی‌های کلاسیک و نظریه بازی‌های تطوری

مؤلفه‌های بازی	نظریه بازی‌های کلاسیک	نظریه بازی‌های تطوری
بازیکنان	انسان عقلایی	موجودات زنده
تعداد بازیکنان مورد تحلیل	به طور معمول اندک	جمعیتی از بازیکنان
استراتژی‌ها	انتخاب‌های ممکن	گونه (سنخ) های ممکن
پیامد	فایده	تناسب زیستی
وضعیت مورد تحلیل	استراتژی تعادلی نش یا مشتقات آن	استراتژی پایدار تطوری

منبع: برگرفته از: (زاهدی وفا و رضائی، ۱۳۹۴، ص ۷۳)

## نظریه بازی‌های کلاسیک و فلسفه اخلاق

غالب نظریه‌پردازان نظریه بازی‌ها نگاهی خنثا به چارچوب تحلیلی نظریه بازی‌ها دارند. به بیان دقیق‌تر، رویکرد آنان راجع به عقلانیت بر تلقی هنجاری از عقلانیت متکی نیست؛ به‌طور مثال، رابرت آومن (Robert Aumann) در تأیید ارزش خنثایی نظریه بازی‌ها باور دارد:

هر چند می‌توان بین نظریه بازی‌ها و اخلاق، ارتباطاتی مفهومی - برای مثال از نوع چگونگی اثرگذاری نظریه بازی‌ها بر اخلاق یا استفاده از نظریه بازی‌ها برای تحلیل رفتارهای اخلاقی نمونه‌هایی - پیدا کرد و از سویی دیگر شناخت این ارتباط به خودی خود اهمیت دارد اما باید خاطرنشان کرد با تمام این اوصاف، نظریه بازی‌ها از هر گونه محتوا و مضمون اخلاقی تهی است؛ به عبارت دیگر، نظریه بازی‌ها هیچ‌گونه توصیه اخلاقی در بر ندارد و از لحاظ اخلاقی به طور کامل خنثاست؛ برای مثال، تعادل نش هیچ‌گاه نمی‌گوید فایده را حداکثر کن بلکه فقط نتیجه‌های آنچه انجام می‌دهیم را کشف می‌کند (Aumann, 2008, p.553).

۱۰۵

کنت بینمور در این‌باره می‌گوید: «نظریه بازی‌ها، رشته‌ای فاقد ماهیت حقیقی و محتوای قائم به ذات است. نظریه بازی‌ها به مردم نمی‌گوید چه باوری داشته یا چه عملی باید انجام دهند...» (Binmore, 2007a, p.56). وی در جای دیگری می‌افزاید:

قضیه‌های ریاضی اساساً این‌همانی‌گویی (Tautology) هستند. این قضیه‌های این‌همانی‌گویانه نمی‌توانند اشتباه باشند چرا که هیچ امر قائم به ذاتی را ادعا نمی‌کنند که بخواهد درست یا غلط باشد. ... گزاره‌های نظریه بازی‌ها نمی‌توانند بد یا خوب باشند. این گزاره‌ها اساساً نسبت به موضوع‌های اخلاقی خنثا هستند (Binmore, 1994, pp.95-96).

آریل رابینشتین (Ariel Rubinstein) نیز در مؤخره‌ای که در شصتمین سالگرد نشر کتاب **نظریه بازی‌ها و رفتار اقتصادی** بر این کتاب می‌نویسد، مدعی هیچ‌گونه دلالت هنجاری برای نظریه بازی‌ها نمی‌شود. وی می‌گوید:

نظریه بازی‌ها به شناسایی الگوهای استدلال پرداخته و دلالت‌های آن را برای تصمیم‌گیری در این وضعیت تعاملی ارزیابی می‌کند؛ بنابراین نظریه بازی‌ها هیچ دلالت هنجاری در خود ندارد ... نظریه بازی‌ها قرابت بسیاری با منطق دارد. منطق، اجازه انتخاب جمله‌های صادق از کاذب را نمی‌دهد، منطق به ما در انتخاب درست از غلط کمک نمی‌کند. نظریه بازی‌ها نیز به ما نمی‌گوید کدام عمل مرجح است (Neumann & Morgenstern, 2007, pp.634-635).

در برابر، برخی نوآوران نظریه بازی‌ها بر وجود تلقی محتوایی از عقلانیت، مبتنی بر پی‌گیری منافع شخصی به عنوان اصل راهنما در موقعیت‌های انتخاب پافشاری دارند. شاید مهم‌ترین تصریح بر جنبه هنجاری بودن نظریه بازی‌ها را بتوان در هارزانی مشاهده کرد؛ وی بر این باور است که در نظریه بازی‌ها مانند غالب نظریه‌های اقتصادی، نظریه ما یک نظریه هنجاری (تجویزی) (Prescriptive) است تا اینکه یک نظریه اثباتی (توصیفی) باشد. دست‌کم آنکه ... آن - نظریه - با پرسش از اینکه هر بازیکن باید به منظور ارتقای مؤثرتر منافع شخصی در بازی چگونه عمل کند، سروکار دارد و نه با پرسش از اینکه وی - یا افرادی مشابهش - چگونه به واقع در یک بازی عمل می‌کنند ( Harsanyi, 2008 p.16, [1977]). به بیان ساده، سرانجام تحلیل‌های نظریه بازی‌ها را می‌توان توصیه راجع به شیوه انتخاب تلقی کرد، هر چند که مدعی شویم الزام به انتخاب استراتژی پیشنهادی از طرف خود نظریه بازی‌ها نیست.

به نظر می‌رسد اگر بخواهیم به آنچه متخصصان نظریه بازی‌ها انجام می‌دهند و نه آنچه متخصصان نظریه بازی‌ها ادعا می‌کنند توجه کنیم، می‌بینیم که نهایت کوشش آنان برای درک موقعیت تعادلی - موقعیتی که انحراف از آن برای بازیکنان مطلوب نیست - صرف می‌شود. \* یافتن این تعادل نیز چه به‌طور صریح و چه ضمنی دلالت بر نوعی توصیه یا تجویز دارد؛ وضعیت مطلوبی که بر پایه مفروض‌های مورد نظر درباره عقلایی بودن بازیکنان به آن خواهیم رسید یا اینکه توصیه آن است که به آن گونه عمل شود. \* به این

---

\*. به قول این بیان متناسب به /ینشتاین که اگر می‌خواهید چیزی از فیزیکدانان درباره روش‌هایی که به کار گرفته‌اند بدانید «به حرف‌های آنها گوش ندهید و توجه خود را معطوف به اعمال آنها کنید».

\*\* همان‌طور که قابل پیش‌بینی است چنین جمع‌بندی موافقان و نیز مخالفان خود را دارد اما نکته‌ای که به نظر می‌رسد تا حد فراوانی بر هنجاری یا توصیه‌ای بودن نظریه بازی‌ها دلالت دارد، اثرگذاری‌های غیرمستقیم یا غیرصریح نظریه بازی‌ها بر رفتار افراد است. در خیل عظیم انواع بازی‌ها و موقعیت‌های تعاملی می‌توان مواردی را یافت که پس از ارائه تحلیل مبتنی بر نظریه بازی‌ها، آزمایش‌شونده به گونه‌ای متفاوت از آنچه پیش از ارائه این تحلیل‌ها رفتار می‌کرد، دست به انتخاب می‌زند؛ شاید بتوان از این پدیده به‌عنوان کاربرد استعاره‌گونه نظریه بازی‌ها یاد کرد: فرد با اطلاع از ساختار بازی‌ها یا موقعیت‌های خاص خواهد کوشید مبتنی بر حسابگری دقیق و عقلایی مد نظر تحلیل‌های نظریه بازی‌ها عمل کند تا دچار ناسازگاری در رفتار خویش نشود.



ترتیب، اساساً مفهوم تعادل در عین برخورداری از ظاهری توصیفی، مقوله‌ای هنجاری بوده و با بحث‌های اخلاق هنجاری قرابت می‌یابد.\*

گفتنی است که پیشنهاد بهره‌گیری از نظریه بازی‌ها در فلسفه اخلاق نیز که نخستین بار به وسیله فیلسوفی ارائه شده، فقط در صورت هنجاری بودن یا بهره‌گیری‌های هنجاری از نظریه بازی‌ها معنادار خواهد بود.

این شکایتی مداوم در برابر فیلسوفان اخلاق دانشگاهی است که در برج عاجی ساکن هستند که در آن و به منظور تحلیل مفهوم‌های اخلاقی یا کاوش معنایی منطقی (Logical Semantics) زبان اخلاقیات به اندیشه فرو رفته‌اند؛ آن هم بدون هبوط به زمین به منظور کاربردی کردن مطالعه‌های انتزاعی‌شان برای اعطای توصیه کاربردی راجع به چگونگی حل مسائل واقعی (Actual) اخلاقی (Braithwaite, 1955, p.3).

این بند، عبارت نخست سخنرانی افتتاحی ریچارد بریث‌ویت (Richard B. Braithwaite) هنگام تصاحب کرسی استادی فلسفه نایتزبریج در دانشگاه کمبریج در سال ۱۹۵۴م است. نطق اثرگذاری که حاوی پیشنهادی برای پژوهشگران حوزه فلسفه اخلاق بود؛ عنوان این سخنرانی به خوبی گویای این توصیه است: «نظریه بازی‌ها به مثابه ابزاری برای فیلسوف اخلاق».

حال که هنجاری بودن نتیجه تحلیل‌های نظریه بازی‌های کلاسیک یا دست‌کم وجود چنین رویکردی در میان برخی اقتصاددانان و فیلسوفان را مفروض پنداشتیم، نوبت به ماهیت این تجویزی

---

\*. همین مطلب راجع به خود علم اقتصاد نیز صادق است. اگر تمام کوشش‌های مطالعه‌های اقتصادی را فقط توصیف تلقی کنیم، با این پرسش روبه‌رو می‌شویم که «سیاست‌گذاری و کنترل پدیده‌های اقتصادی بر چه اساس و معیاری استوار می‌شود؟ اگر قرار باشد دو وضعیت اقتصادی با هم مقایسه و یکی انتخاب شود، چگونه باید این انتخاب صورت گیرد؟» در حقیقت - در صورت نفی تجویزی بودن علم اقتصاد - برای وفادار ماندن به اثباتی بودن گزاره‌ها، این حوزه از دانش را بی‌ثمر کرده‌ایم (عربی، ۱۳۸۷، ص ۴۳). هر چند جای توضیح دقیق‌تر مجموعه کاربردهای نظریه بازی‌ها نیست اما دستاوردهای حوزه‌هایی مانند طراحی سازوکار یا طراحی بازار به عنوان دلالت‌های مهندسی‌گونه نظریه بازی‌ها، تمام کوشش خویش را برای تغییر جهان و دست‌یافتن به وضعیت مطلوب، آن هم از دید معیارهای مد نظر طراح سازوکار، مصروف داشته و اساساً نشانه‌ای آشکار از تجویزی یا هنجاری بودن هستند.

یا هنجاری بودن می‌رسد. با نگاهی اجمالی به شیوه تحلیل در نظریه بازی‌های کلاسیک به نظر می‌رسد نظریه بازی‌ها چیزی جز همان ایده‌های نتیجه‌گرایانه خودگروانه (Egoistic)\* نیست.

همان‌طور که اشاره شد، نظریه بازی‌های کلاسیک، مطالعه تعامل‌های بازیکنان عقلایی است. به محض برخوردن با مفهوم عقلانیت باید تأملی بر دخالت این مفهوم بر چارچوب تحلیلی نظریه بازی‌ها کرد. اساس یافتن تعادل‌های مختلط نش بر نظریه مطلوبیت انتظاری سامان یافته است. لازمه استفاده از نظریه مطلوبیت انتظاری و دست‌زدن به غالب تحلیل‌های نظریه بازی‌ها نیز آن است که فرد بتواند برای تمام گزینه‌های پیش روی خویش، مقداری عددی به‌عنوان شاخص ارزشمندی آن گزینه برگزیند؛ به‌طور مثال، در یک بازی ساده مانند معمای زندانی باید با لحاظ‌کردن تمام ویژگی‌های شخصیتی، احساسات، منافع و مضار یا باور خویش راجع به ارزش‌مندی انتخاب‌های ممکن، مقدار عددی برای استراتژی همکاری و نیز برای استراتژی عدم همکاری برگزیده شود. ضرورت تعیین پیامد انتخاب‌های ممکن، بیانگر ویژگی مهمی در بهره‌گیری‌های اخلاقی از نظریه بازی‌هاست. این نظریه فقط می‌تواند در خدمت نظریه‌های اخلاقی نتیجه‌گرا باشد.

نمونه بارز چنین محدودیتی را می‌توان در ترجمه‌ناپذیر بودن رفتارهای مبتنی بر نظریه‌های رقیب اخلاقی در نظریه بازی‌ها دانست. به نظر می‌رسد بهترین اشاره به این کاستی، پیش از این به‌وسیله توماس شلینگ (Thomas Schelling, 1968) صورت گرفته که عیناً به آن اشاره می‌کنیم. پرسش مهمی که از نظر وی در ربط میان نظریه بازی‌ها و اخلاق امکان بروز دارد آن است که آیا روش نظریه بازی‌ها در مقام کاربرد به طبقه خاصی از اخلاق محدود است؟ وی از مثالی بهره می‌برد که می‌تواند قابل تطبیق با اخلاق تکلیف‌محور باشد. وی فردی را مثال می‌زند که تمام علاقه و هدفش انجام انتخاب حق است: وی علاقه‌ای به نتیجه‌ها ندارد، وی فقط می‌خواهد کار درست را انجام دهد. اینجاست که تا حد فراوانی می‌توان با شلینگ هم‌باور بود که نظریه بازی‌ها نمی‌تواند چندان کمکی کند؛ آن - یعنی نظریه بازی‌ها، فقط - با رابطه نتیجه‌های (Consequences) انتخاب‌ها سروکار دارد (Schelling, 1968, p.42).

\*. اینکه شیوه تحلیل در نظریه بازی‌های کلاسیک مبتنی بر نتیجه است تا حدی واضح به نظر می‌رسد اما اینکه این شیوه تحلیل با چه نوعی از نتیجه‌گرایی (خودگروی، فایده‌گرایی عمل‌محور، فایده‌گرایی قاعده‌محور یا ...) سازگار است می‌تواند محل تأمل بیشتری قرار گیرد.

در صورت بهره‌گیری از چارچوب تحلیلی نظریه بازی‌ها، نادیده گرفته شدن مؤلفه اخلاقی دیگری نظیر «نیت» نیز به راحتی قابل مشاهده است. اساس تحلیل‌های نظریه بازی‌ها بر اساس اطلاعات موجود از ساختار بازی صورت می‌گیرد؛ اطلاعاتی مانند پیامدهای وضعیت‌های گوناگون و مجموعه انتخاب‌های ممکن بازیکن رقیب. در مجموعه اطلاعاتی که بر اساس آن به تحلیل و مطالعه انتخاب پرداخته می‌شود، عنصر نیت وجود نداشته و انتخاب‌های افراد رقیب مبتنی بر فرض پی‌گیری منافع شخصی (عقلانیت) انگاشته می‌شود.

برای تبیین مناسب‌تر مطلب پیشین، ارائه مثالی از شلینگ (Ibid, p.43) مفید به نظر می‌رسد. فرض کنید یک بازی ساده مانند معمای زندانی یا شکار گوزن، میان دو بازیکن بتواند تا بی‌نهایت بار تکرار شود (بازی تکراری). یکی از استراتژی‌های تعادلی در این بازی آن است که بر اساس قاعده زیر عمل شود: «با رقیب همکاری کن تا زمانی که وی استراتژی عدم همکاری را برگزیند، در این صورت تا ابد استراتژی عدم همکاری را برگزین». در این صورت اگر بر فرض شما در مرحله  $n$  ام به هر علتی بر خلاف مرحله‌های پیشین قبل، استراتژی عدم همکاری را انتخاب کردی، بازیکن رقیب بر اساس قاعده راهنمای استراتژی تعادلی، در مرحله  $n+1$  عدم همکاری را برمی‌گزیند. بر همین اساس شما نیز در مرحله  $n+2$  و مرحله‌های بعدی استراتژی عدم همکاری را انتخاب می‌کنید؛ اما نکته ظریفی در اینجا نهفته است: می‌توان دو هدف یا نیت برای عدم همکاری فرد رقیب در مرحله  $n+1$  متصور بود: الف) رقیب به آن علت عدم همکاری را برگزید که شما انتخابی برخلاف حداکثر شدن منافع وی در مرحله  $n$  انجام داده‌اید؛ ب) رقیب به علت تعهد و پای‌بندی به اصلی اخلاقی و با نیت مجازات فردی که نوعی توافق ضمنی را زیر پا گذاشته، عدم همکاری در مرحله  $n+1$  را انتخاب کرده است.

اگر نیت بازیکن رقیب، نیت نخست باشد، عدم همکاری در مرحله  $n+2$  می‌تواند معنادار باشد، چون وی بر پایه کاهش منافعش شما را مجازات کرده و شما نیز به عدم همکاری اقدام می‌کنید؛ اما اگر هدف رقیب از عدم همکاری در مرحله  $n+1$  نیت دوم باشد، آیا نباید به واسطه این نیت اخلاقی و تعهد وی به آن، در مرحله  $n+2$  استراتژی همکاری را برگزینیم تا وی را به واسطه این نیت منتفع سازیم؟ نکته اساسی اینجاست که انواع این نیت‌ها در نظریه بازی‌ها مورد سنجش یا نظر قرار نمی‌گیرند. به بیان دقیق‌تر و بر اساس مفروض‌های نظریه

بازی‌ها، نیت قابل ورود به تحلیل نبوده و اساساً بر اساس فرض عقلانیت و وجود دانش مشترک راجع به این فرض در میان بازیکنان، تحلیل‌های نظریه بازی‌های کلاسیک شکل می‌گیرد. حال که فقط بر اساس پیامدها و انتخاب‌هاست که تحلیل‌ها صورت می‌گیرد و باقی مؤلفه‌ها نادیده گرفته می‌شوند، روشن خواهد بود که مؤلفه‌ای که از نظر برخی مکتب‌های اخلاقی و حتی درک عرفی از اخلاق باید مورد نظر باشد، حذف می‌شوند.

البته ممکن است در مقام دفاع از نظریه بازی‌ها این گونه برآمد که قرار نیست این نظریه توانایی در نظرگرفتن تمام ارکان یا عناصر مربوط به نظریه‌های اخلاقی گوناگون را داشته باشد. در برابر و ابتدا باید اعلام داشت که برخی بر این باور هستند که این نظریه، توانایی تحلیل انواع متنوعی از مؤلفه‌های رفتاری انسان مانند نتیجه عمل، انگیزه فاعل یا رفتارهای مبتنی بر تکلیف‌گرایی را دارد (Hoffman, Yoeli & Navarrete, 2016).

دشواری نخست از آنجا آشکار می‌شود که اساس شیوه تحلیل در نظریه بازی‌ها شهود (Perception) متخصص نظریه بازی‌هاست. به بیان رابینشتین، نظریه بازی‌ها «مدل‌ها را بر اساس شهود - فرد به کارگیرنده آن - می‌سازد (Rubinstein, 1991, p.923)». از این دید ارتباط «نظریه بازی‌ها» با «واقعیت» به معنای تناظر تحلیل‌هایش با آنچه در دنیای خارج روی می‌دهد نیست بلکه بیشتر از حیث «شهود ما» از «پدیده‌های واقعی» است (Ibid, p.909). اینجاست که دشواری نخست پیش روی انگاره «توانایی تحلیل تمام رفتارهای تعاملی» ظاهر می‌شود.

گذشته از انتقاد پیشین و حتی با فرض انطباق شهود به کارگیرنده نظریه بازی‌ها با واقعیت، بحث مهم دیگری در می‌گیرد و آن امکان‌پذیر بودن ترجمه این شهود منطبق با واقع در قالب چارچوب‌های تحلیلی نظریه بازی‌هاست؛ دشواری عمده در سامان‌دادن جدول پیامدها برای بازیکنان است. به نظر می‌رسد که تحلیل‌های نظریه بازی‌ها، به نحوی از انحاء به تحلیل‌های نتیجه‌گرایانه مقید است. واضح خواهد بود که به واسطه نتیجه‌گرایانه یا فایده‌گرابودن چارچوب تحلیلی در نظریه بازی‌های کلاسیک، دشواری‌های پیش روی این نظریه اخلاقی از جمله سنجش ناپذیری فایده و عدم امکان مقایسه بین شخصی فایده، برای آن هم وجود دارد که البته در اینجا به آن نمی‌پردازیم. نتیجه آنکه با پی‌بردن به گستره کاربردپذیری نظریه بازی‌های کلاسیک می‌توان اعلام داشت که تمام مسائل یا دوراهی‌های اخلاقی را نمی‌توان به سادگی در قالب مدل‌های نظریه بازی‌ها ترجمه کرد؛ بر خلاف تلقی‌ای که بریث ویت از این نظریه برای حل مسائل واقعی اخلاقی داشت.

## نظریه بازی‌های تطوری و فلسفه اخلاق

با توجه به تفکیکی که در قالب جدول یک میان نظریه بازی‌های کلاسیک و نظریه بازی‌های تطوری ارائه شد، واضح خواهد بود که در اینجا اساساً مسئله انتخاب انسان به کناری گذاشته شده و به جای آن طبیعت، برآمد تعامل‌های جمعیتی افراد، قدرت تولید و اشاعه نوع خویش یا تناسب زیستی اثرگذار خواهد بود. در این صورت می‌توان ابتداء این نظریه را بر ایده‌های داروینیستی دانسته؛ در نتیجه، کاربردهای آن در حوزه مقوله‌های اخلاقی مانند احساس اخلاقی، عدالت، انصاف و حق را نیز ترجمان اخلاق طبیعت‌گرایانه و اخلاق تطوری (Evolutionary Ethics) برشمرد که با پیشگامی چارلز داروین (Charles Darwin) و هربرت اسپنسر (Herbert Spencer) مطرح شده است.

در این حوزه از نظریه بازی‌ها نیز قیدهایی مربوط به مفروض‌ها و چارچوب تحلیلی مسببات محدودشدن دامنه بهره‌گیری از نظریه بازی‌های تطوری در بحث‌های فلسفه اخلاق را فراهم خواهد کرد.

یکی از تفاوت‌های مبانی اخلاقی این دسته از نظریه‌ها با نظریه بازی‌های کلاسیک به بُعد توصیفی آن بر می‌گردد. اگر نظریه بازی‌های کلاسیک را نوعی تجویز یا توصیه بدانیم، نظریه بازی‌های تطوری با توجه به مفروض‌هایش، منطقی‌تر در حوزه توصیف بوده و نمی‌تواند ارائه‌کننده اصل راهنمای عملی در موقعیت‌های مربوط به دوراهی‌های اخلاقی باشد. اساساً کاربردی‌سازی نظریه بازی‌های تطوری در موضوع‌هایی مانند مالکیت، عدالت، انصاف یا حق، نه به قصد دست‌یابی به شیوه رسیدن به موقعیت عادلانه بلکه فقط برای تحلیل آنچه روی داده کاربرد دارد؛ تمام تحلیل‌های مبتنی بر این نظریه مانند خود نظریه داروین، فاقد ویژگی پیش‌بینی‌پذیری یا توصیه‌مندی است. نتیجه‌ها و دلالت‌های مطالعه‌های ساگدن و اسکرمز که به آنها اشاره‌ای صورت گرفت، به‌خوبی نشان‌دهنده این امر است.\*

---

\*. تأکید می‌شود که منظور از هنجاری‌بودن در اینجا آن است که اصل یا اصول راهنمای عمل برای افراد ارائه شود. اصل بقای انسب راهکاری برای ما جهت انتخاب یا گزینش ارائه نمی‌کند. در مقام تشبیه، ایده داروین راهکاری به ما انسان‌ها برای مقطوع‌النسل کردن برخی گونه‌های جانوری یا توجه بیش از حد به گونه‌های دیگر ارائه نمی‌کند. از درون ایده داروین یا اسپنسر یا موافقان داروینیسم اجتماعی توصیه‌ای برای حذف یک قانون، سنت یا اشاعه یک نهاد استخراج نمی‌شود.

به‌طور خلاصه نظریه بازی‌های تطوری به انضمام فروض و چارچوب تحلیلی محدودکننده‌اش، اساساً فاقد شرط اولیه برای دستیابی به هدف بریث ویت یعنی ارائه توصیه در مسائل اخلاقی واقعی بوده و در حقیقت فقط ابزاری برای طرح علمی - ریاضی ایده داروین و اسپنسر است.

تمسک به چنین تلقی از مفهوم‌هایی مانند اخلاق، عدالت یا انصاف جایی برای بسیاری از موضوع‌های مورد نظر در فلسفه اخلاق نمی‌گذارد؛ روایی یا ناروایی اعمال، خوبی یا بدی نتیجه‌ها، فضیلت‌مندی یا رذیلت‌مندی ویژگی‌های شخصی یا خیر و شر بودن محرک‌ها و اراده عمل اساساً در این تحلیل‌ها وجود نداشته و مطالعه پژوهشگر اخلاق به تبیین رویدادهای گذشته آن هم در مقام توصیف شیوه سامان‌یافتن نهادها، شکل‌گیری عادت‌های فردی، میثاق‌های اجتماعی و اثر آنها بر انتخاب افراد یا اجتماع‌ها منحصر می‌شود.

### جمع‌بندی و نتیجه‌گیری

توجه به حوزه مطالعاتی فلسفه اخلاق و اهمیت مقوله عقلانیت و نیز بروز انواع موقعیت‌های تعاملی در برخی مسائل اخلاقی، ما را به امکان‌پذیر بودن اثرپذیری مطالعه‌های فلسفه اخلاق از نظریه بازی‌ها ظنین می‌کند. مقاله پیش رو در وهله نخست، اثرگذاری برخی ایده‌ها و مدل‌های نظری در نظریه بازی‌ها بر برخی نظریه‌ها در حوزه اخلاق و عدالت را نشان داد. در یافتیم که به‌طور صریح یا ضمنی، برخی ایده‌ها در حوزه اخلاق از ظهور نوآوری‌ها در نظریه بازی‌ها متأثر شده و فرضیه نخست مقاله تأیید شد.

با ارزیابی چنین ارتباطی، در یافتیم که مجموعه مطالعه‌های نظریه بازی‌های کلاسیک با نوع خاصی از انواع نظریه‌های اخلاقی سازگار بوده و برای طرح بحث از مقوله‌ها و ایده‌هایی مانند تکلیف‌گرایی، نیت فاعل یا فضیلت‌مداری آماده و مهیا نیست. همچنین در استفاده‌هایی که از نظریه بازی‌های تطوری در موضوع‌های اخلاقی شده است، در یافتیم که به‌واسطه محدودیت‌ها در مفروض‌ها و چارچوب تحلیلی‌اش، اساساً از حوزه مطالعه‌های فلسفه اخلاق و بحث‌هایی مانند روایی یا ناروایی اعمال، خوبی یا بدی نتیجه‌ها، فضیلت‌مندی یا رذیلت‌مندی ویژگی‌های شخصی یا خیر و شر بودن محرک‌ها و اراده عمل خارج می‌شویم. در نتیجه فرضیه دوم مقاله به این پاسخ رهنمون می‌شود که چارچوب تحلیلی نظریه

بازی‌ها نمی‌تواند به‌عنوان ابزاری کارا برای فلسفه اخلاق به طوری که امکان بهره‌مندی از آن در تمام یا غالب نظریه‌های اخلاقی وجود داشته باشد، استفاده شود؛ اولاً، از آن جهت که ایده‌های نظریه بازی‌ها تناسب بیشتری با برخی نظریه‌های اخلاقی خاص دارد؛ ثانیاً، به این علت که متخصص نظریه بازی‌ها در تحلیل و تفسیر موقعیت‌های تعاملی نقش بسیار مهمی ایفا می‌کند؛ به بیان دیگر، این شهود متخصص نظریه بازی‌ها به‌عنوان مشاهده‌گر است که از بین عنوان‌ها و مفهوم‌های گوناگونی مانند پی‌گیری منافع شخصی، پی‌گیری بیشترین منفعت برای بیشترین افراد، عمل بر اساس تکلیف یا توجه به نیت دست به انتخاب زده و اساس تحلیل خود را بر آن استوار می‌کند.

## منابع و مأخذ

۱. پالمر، مایکل؛ **مسائل اخلاقی**: متن آموزشی فلسفه اخلاق؛ تهران: سازمان مطالعه و تدوین کتب علوم انسانی دانشگاه‌ها و پژوهشگاه علوم و فرهنگ اسلامی، ۱۳۸۸.
۲. پیک حرفه، شیرزاد؛ «سودای سربالا: سقف اخلاق بر ستون سود»؛ پژوهش‌های فلسفی، س ۵، ش ۹، ۱۳۹۰.
۳. زاهدی‌وفا، محمدهادی و محمدجواد رضایی؛ «نظریه بازی‌های تکاملی و ارزیابی مجدد از ایده «پیشرفت در نظریه بازی‌ها»؛ بررسی از نظر روش‌شناسی علم اقتصاد»؛ فلسفه علم، س ۵، ش ۲، ۱۳۹۴.
۴. عربی، سیدهادی؛ «از اقتصاد اثباتی تا اقتصاد هنجاری: بنیان‌های فلسفی و ارزشی اقتصاد رفاه»؛ روش‌شناسی علوم انسانی، س ۱۴، ش ۵۷، ۱۳۸۷.
۵. مصباح یزدی، محمدتقی؛ نقد و بررسی مکاتب اخلاقی؛ قم: مؤسسه آموزشی و پژوهشی امام خمینی علیه السلام، ۱۳۸۴.
6. Aumann, R. J.; "Game Theory"; in **The New Palgrave Dictionary of Economics**, ed. by Steven N. Durlauf and Lawrence E. Blume, Basingstoke: Palgrave Macmillan, 2008.
7. Binmore, Kenneth G.; **Playing fair: Game Theory and the Social Contract**; Cambridge, Mass., London: The MIT Press, 1994.

8. —; **Just playing: Game Theory and Social Contract**; Cambridge, Mass., London: MIT Press, 1998.
9. — ; **Natural justice**; New York, Oxford: Oxford University Press, 2005.
10. —; **Game theory: A very short introduction**; New York: Oxford University Press, 2007a.
11. —; **Playing for Real: A text on game theory**; Oxford: Oxford University Press, 2007b.
12. Braithwaite, Richard B.; **Theory of games as a tool for the moral philosopher: An inaugural lecture delivered in Cambridge on 2 December 1954**; Cambridge: Cambridge University Press, 1955.
13. Curry, Oliver S.; “Morality as Cooperation: A Problem-Centred Approach”; in **The evolution of morality**; ed. by Todd K. Shackelford and Ranald D. Hansen; Cham, Switzerland: Springer, Evolutionary psychology, 2016.
14. Dixit, Avinash K., and Susan Skeath; **Games of strategy**; 2nd ed; New York, London: W. W. Norton & Co, 2004.
15. Gauthier, David P; **Morals by agreement**; Oxford: Clarendon, 1986.
16. —; “Uniting separate persons”, in **Rationality, justice and the social contract: Themes from Morals by agreement**; ed. by David P. Gauthier and Robert Sugde; New York, London: Harvester Wheatsheaf, 1993.
17. Gibbons, Robert; **A primer in game theory**; Harvester Wheatsheaf, 1992.



18. Harsanyi, John C.; "Can the Maximin Principle Serve as a Basis for Morality? A Critique of John Rawls's Theory"; **The American Political Science Review**, 69.2, <http://www.jstor.org/stable/1959090>, 1975.
19. —; **Rational behavior and bargaining equilibrium in games and social situations**; 1st ed; New York: Cambridge University Press, Cambridge paperback library, 2008 [1977].
20. Hausman, Daniel M., and Michael S. McPherson; **Economic analysis, moral philosophy, and public policy**; 2nd ed; New York: Cambridge University Press, 2006.
21. Hoffman, Moshe, Erez Yoeli, and Carlos D. Navarrete; "Game Theory and Morality", in **The Evolution of Morality**; ed. by K. T. Shackelford and D. R. Hansen; Cham: Springer International Publishing, 2016.
22. Kalai, Ehud, and Meir Smorodinsky; "Other Solutions to Nash's Bargaining Problem", **Econometrica**, 43.3, 1975.
23. Lehtinen, A., and J. Kuorikoski; "Unrealistic Assumptions in Rational Choice Theory"; **Philosophy of the Social Sciences**, 37.2, 2007.
24. Nash, John F; "The Bargaining Problem"; **Econometrica**, 18.2, <http://www.jstor.org/stable/1907266> ,1950.
25. Neumann, John von, and Oskar Morgenstern; **Theory of games and economic behavior**; 60th ed; Princeton: Princeton University Press, Princeton classic editions; 2007.
26. Pettit, P.; "Virtus normativa: Rational choice perspectives"; **Ethics**, 1990.
27. Rasmusen, Eric; **Games and Information: Reputation and Repeated Games with Symmetric Information**; [http://www.rasmusen.org/GI/chapters/chap05\\_repeated.pdf](http://www.rasmusen.org/GI/chapters/chap05_repeated.pdf), 2005.

28. Rawls, John; **A Theory of Justice**; Cambridge, MA: Belknap Press of Harvard University Press, 2005
29. Rubinstein, Ariel; “Comments on the Interpretation of Game Theory”; **Econometrica**, 59.4, <http://www.jstor.org/stable/2938166>, 1991.
30. Schelling, Thomas C.; “Game Theory and the Study of Ethical Systems”; **The Journal of Conflict Resolution**, 12.1: 34–44 <http://www.jstor.org/stable/172812>, 1968.
31. Skyrms, Brian; **Evolution of the social contract**; Cambridge: Cambridge University Press. 1996.
32. —; **The Stag Hunt and the Evolution of Social Structure**; Cambridge: Cambridge University Press, 2004.
33. Smith, John M.; ‘Game theory and the evolution of fighting’, in **On evolution**; ed. by John M. Smith; Edinburgh: Edinburgh University Press, 1972.
34. Smith, John M., and G. R. Price; “The Logic of Animal Conflict”; **Nature**, 246.5427, 1973.
35. Sugden, Robert; “Spontaneous Order”; **The Journal of Economic Perspectives**, 3.4, 1989.
36. —; **The economics of rights, co-operation, and welfare**; 2nd ed; New York: Palgrave Macmillan, 2005 [1986].
37. Verbeek, B.; “Rational Choice Virtues”; **Ethical theory and moral practice**, 13(5), 2010.