

# Theory of Mind: A Review of Recent Empirical Results and Main Theoretical Explanations

---

**Legg, Edward W.**

*Source / Izvornik:* **Psihologijske teme, 2023, 32, 635 - 661**

**Journal article, Published version**

**Rad u časopisu, Objavljena verzija rada (izdavačev PDF)**

<https://doi.org/10.31820/pt.32.3.12>

*Permanent link / Trajna poveznica:* <https://um.nsk.hr/um:nbn:hr:186:446669>

*Rights / Prava:* [In copyright](#)/[Zaštićeno autorskim pravom.](#)

*Download date / Datum preuzimanja:* **2024-06-26**



*Repository / Repozitorij:*

[Repository of the University of Rijeka, Faculty of Humanities and Social Sciences - FHSSRI Repository](#)



## Teorija uma: pregled suvremenih empirijskih rezultata i glavnih teorijskih objašnjenja

Edward William Legg

Sveučilište u Rijeci, Filozofski fakultet, Odsjek za psihologiju, Rijeka, Hrvatska

---

### Sažetak

Teorija uma je sposobnost pripisivanja mentalnih stanja drugima. Do prije otprilike 20 godina empirijska istraživanja upućivala su na to da je teorija uma kognitivno zahtjevna sposobnost koja se razvija oko četvrte godine života. Međutim, sve veći broj istraživanja koja koriste zadatke koji ne zahtijevaju verbalne odgovore upućuje na to da sposobnost koja nalikuje na teoriju uma postoji i izvan kognitivne kontrole te da je pokazuju već i djeca stara devet mjeseci. Ovaj pregledni rad predstaviti će tri glavna tipa teorija koje pokušavaju objasniti te novije empirijske rezultate i raspraviti o njima. Nadalje, u radu će biti istaknuti empirijski nalazi koji idu u prilog svakoj od tih teorija ili je opovrgavaju. Iako postoje prijedlozi za kritične eksperimente koji bi trebali moći napraviti razliku između teorija, trenutačno ti prijedlozi još nisu dovoljno korisni za te svrhe jer se podatci koji bi se mogli dobiti takvim eksperimentima mogu u različitim teorijama različito interpretirati. Stoga je važno da se budući rad o teoriji uma fokusira na uklanjanje dvosmislenosti u predviđanjima i interpretacijama svake od teorija.

*Ključne riječi:* teorija uma, socijalna kognicija, zauzimanje perspektive

---

### Uvod

Pojam *teorija uma* skovali su Premack i Woodruff 1978. godine, a on označava sposobnost reprezentacije i rasuđivanja o mentalnim stanjima (Premack i Woodruff, 1978). Premack i Woodruff razradili su pojam teorije uma na temelju rezultata istraživanja s jednom čimpanzom, međutim, istraživanja o toj temi promijenila su se od 1978. godine te su zahvatila i područja izvan komparativne psihologije, ponajprije razvojnu psihologiju, neuroznanost, psihologiju starosti te klinička istraživanja (Frith i Frith, 2005; Saxe i Baron-Cohen, 2006).

Većina istraživanja tijekom prvih 20 godina nakon početka istraživanja teorije uma ukazuje na to da je riječ o kognitivno zahtjevnoj sposobnosti koja se vrlo vjerojatno kasno razvija te koju možda posjeduju samo ljudi. Primjerice, podatci iz

---

Edward William Legg  <https://orcid.org/0000-0003-0353-7105>

✉ Edward William Legg, Odsjek za psihologiju, Filozofski fakultet u Rijeci, Sveučilišna avenija 4, 51000 Rijeka, Hrvatska. E-adresa: [edward.legg@uniri.hr](mailto:edward.legg@uniri.hr)

razvojne psihologije ukazuju na to da djeca mlađa od četiri godine ne mogu predvidjeti radnje osoba s pogrešnim vjerovanjem (za pregled v. Wellman i sur., 2001). Istraživanja iz područja socijalne psihologije i komunikacije pokazala su da je i za djecu i za odrasle rasuđivanje o vjerovanjima i perspektivi drugih osoba kognitivno zahtjevan i spor proces koji je sklon pogreškama (Keysar i sur., 1998; Royzman i sur., 2003).

Međutim, u posljednjih 20 godina razna bihevioralna istraživanja dovela su u pitanje to „tradicionalno” gledište teorije uma. Eksperimentalna istraživanja u kojima i djeca od devet mjeseci mogu svoje ponašanje uskladiti s vjerovanjem druge osobe osporavaju hipotezu o kasnome razvoju teorije uma (Kovács i sur., 2010; Onishi i Baillargeon, 2005; Southgate i sur., 2007). Nadalje, gledište prema kojoj je teorija uma sposobnost koja je spora i pod kognitivnom kontrolom pojedinca dovedeno je u pitanje podacima koji ukazuju na to da na prosudbe odraslih osoba može automatski utjecati ono što vidi ili vjeruje druga osoba koja je prikazana u zadacima, čak i u onim situacijama u kojima mentalna stanja druge osobe nisu izravno relevantna za samu prosudbu (Schneider i sur., 2017).

Navedena nova empirijska istraživanja zahtijevala su nove teorije koje bi trebale moći objasniti i one nalaze koji su u skladu s tradicionalnim gledištem, kao i one koji nisu u skladu s njime. Općenito, te teorije možemo podijeliti u tri kategorije: i) *teorije koje su utemeljene na jednome sustavu* koje tvrde da se i tradicionalni i novi podatci mogu objasniti istom temeljnom sposobnošću teorije uma; ii) *nementalističke teorije* koje predlažu da se novi podatci mogu objasniti procesima koji nisu teorija uma; i iii) *teorije koje su utemeljene na dvama sustavima* koje predlažu da postoje dva oblika teorije uma – jedan sustav koji je u skladu s tradicionalnim gledištem teorije uma, tj. kognitivno zahtjevan sustav koji je sposoban prikazati sve značajke mentalnih stanja, i drugi sustav koji nije kognitivno zahtjevan, ali u kojemu su mentalna stanja reprezentirana uz određena ograničenja.

Prvi dio ovoga rada donosi pregled istraživanja teorije uma predstavljajući nalaze koji su doveli do tradicionalnoga gledišta, kao i one nalaze koji predstavljaju izazov tomu gledištu. U drugome se dijelu rada raspravlja o različitim teorijama utemeljenima na novim istraživanjima. Na kraju rada nalazi se pregled empirijskih istraživanja koja testiraju navedene teorije. Potrebno je istaknuti da je fokus ovoga preglednog rada na istraživanjima o teoriji uma kod ljudi s tipičnim razvojem jer je najdetaljnija debata o teorijskim objašnjenjima u literaturi o teoriji uma prisutna upravo u teorijskim i empirijskim radovima o ljudima s tipičnim razvojem.

## Pregled istraživanja teorije uma

### Tradicionalno gledište teorije uma

Rad Premacka i Woodruffa (1978) potaknuo je raspravu o tome što bi bio idealan test za teoriju uma. Mentalna stanja definiraju njihova značajka da posjeduju neki sadržaj – drugim riječima, ona su *o nečemu*. Međutim, sadržaj mentalnih stanja ne mora izravno odgovarati stanju stvari u stvarnosti u tome trenutku, već umjesto toga može predstavljati pogrešan, zastarjeli ili hipotetski sadržaj (Pylyshyn, 1978). Ta je karakteristika mentalnih stanja ključna u prijedlozima testova za teoriju uma. Naime, predlaže se da se u testovima za teoriju uma ispitanicima predstave situacije u kojima su njihova vlastita mentalna stanja (koja i odražavaju stvarnost) u sukobu s mentalnim stanjima o kojima rasuđuju, tj. s mentalnim stanjima neke druge osobe (Bennett, 1978; Dennett, 1978; Harman, 1978), da bi se moglo zaključiti da ispitanici rasuđuju na temelju reprezentacije mentalnih stanje druge osobe, a ne na temelju vlastitoga stanja ili stanja stvari u stvarnosti u tome trenutku. Stoga se smatra da je najbolji test za teoriju uma situacija u kojoj neko mentalno stanje dviju osoba ima različit sadržaj.

Da bi zadovoljili taj kriterij, Wimmer i Perner (1983) razvili su niz zadataka koji ispituju rasuđivanje o pogrešnim vjerovanjima. U tim zadacima djeca u dobi od tri do sedam godina rasuđuju o ponašanju druge osobe kada je vjerovanje te osobe zastarjelo. Zadatci uključuju prikaz osobe nazvane Maxi koja najprije stavlja komadić čokolade u jedan ormarić, a potom napušta scenu pa ne vidi da je jedan od roditelja premjestio čokoladu u drugi ormarić. Wimmer i Perner (1983) otkrili su da djeca mlađa od četiri godine netočno odgovaraju na pitanja povezana s Maxijevim pogrešnim vjerovanjem. To je otkriće potkrijepljeno podacima iz mnogobrojnih verbalnih zadataka pogrešnoga vjerovanja koji pokazuju da ih djeca obično uspješno rješavaju tek nakon četvrte godine (Hogrefe i sur., 1986; Wellman i sur., 2001; Wellman i Liu, 2004).

Neovisno o istraživanjima u razvojnoj psihologiji, istraživanja koja ispituju sposobnost prosuđivanja tuđih unutarnjih stanja i dispozicija, uključujući mentalna stanja, kod odraslih su osoba otkrila pristranost prema vlastitome stanju (Epley i sur., 2004; Gilovich i sur., 1999; Keysar i sur., 1998, 2000, 2003; Nickerson, 1999). Takve se pogreške pristranosti prema vlastitome stanju obično povećavaju kada smo pod kognitivnim opterećenjem, što upućuje na to da je proces prosuđivanja tuđih uvjerenja i perspektiva kognitivno zahtjevan (Lin i sur., 2010).

### *Izazovi tradicionalnomu gledištu teorije uma*

**Podatci iz istraživanja s dojenčadi.** Važan izazov za tradicionalnu konceptualizaciju teorije uma predstavljaju istraživanja koja koriste neverbalne mjere teorije uma. U literaturi o kognitivnome razvoju tu je pretežno riječ o zadacima u kojima se mjeri gdje ispitanik gleda. Clements i Perner (1994) zabilježili

su i verbalne odgovore i mjesto u koje su djeca gledala nakon što su upitana gdje bi osoba prikazana u zadatku s pogrešnim vjerovanjem o lokaciji predmeta tražila taj predmet. U skladu s prijašnjim istraživanjima, Clements i Perner otkrili su da su djeca starija od dvije godine i jedanaest mjeseci, ali mlađa od četiri godine davala pogrešne verbalne odgovore o pogrešnome vjerovanju prikazane osobe. Međutim, ta su ista djeca gledala prema mjestu gdje je osoba pogrešno vjerovala da se predmet nalazi, što znači da su, unatoč netočnim verbalnim odgovorima, pogledima ukazali na to da su na nekoj razini možda točno reprezentirali sadržaj mentalnoga stanja osobe. Na te se rezultate nadograđuju i istraživanja u kojima nisu korištena verbalna pitanja eksperimentatora. Onishi i Baillargeon (2005) upotrijebili su paradigmu kršenja očekivanja koja je utemeljena na tendenciji dojenčadi da dulje gleda u događaje koji su nepoznati ili neočekivani. U tome je istraživanju 15-mjesečna dojenčad dulje gledala u prikaz u kojemu osoba traži predmet na lokaciji koja nije u skladu s njezinim pogrešnim vjerovanjem nego u prikaz u kojemu osoba traži predmet na lokaciji koja je u skladu s njezinim pogrešnim vjerovanjem. Ti rezultati upućuju na to da je dojenčad formirala očekivanje da će se osoba ponašati u skladu sa svojim vjerovanjem. Southgate i suradnici (2007) u svojem su istraživanju upotrijebili sklonost djece da gledaju prema mjestu na kojemu očekuju da će se dogoditi neka radnja (*anticipativno gledanje*, engl. *anticipatory looking*) da bi pokazali da djeca u dobi od 25 mjeseci predviđaju da će osoba s pogrešnim vjerovanjem predmet tražiti na mjestu na kojemu vjeruje da je predmet skriven. Kovács i suradnici (2010) razvili su još jednu paradigmu u kojoj se mjeri koliko dugo dojenčad gleda u događaj koji je u skladu s njihovim vjerovanjem: dojenčad najprije gleda kako se lopta kotrlja iza pregrade, potom lopta izlazi iza pregrade, a nakon što se pregrada spusti, iza nje se, što bi dojenčad i očekivala, ne nalazi lopta. U dvama je uvjetima manipulirano vjerovanjem osobe prikazane na sceni: u jednome je uvjetu osoba bila prisutna na sceni i kada se lopta kotrljala iza pregrade i kada je izišla iza nje (osoba ima isti sadržaj vjerovanja kao i dojenčad), dok je u drugome uvjetu osoba bila prisutna na sceni samo kada se lopta kotrljala iza pregrade, ali ne i kada je lopta izišla iza pregrade (osoba i dojenčad imaju različit sadržaj vjerovanja). Dojenčad stara 7 mjeseci dulje je gledala u događaj kada on nije bio u skladu s vjerovanjem osobe nego kada je događaj bio u skladu s vjerovanjem osobe, iako je događaj u obama uvjetima bio u skladu sa samim vjerovanjem dojenčadi. Ti rezultati upućuju na to da na vlastita očekivanja dojenčadi o tome što će se dogoditi utječe to dijeli li druga osoba njihova očekivanja. Objedinjeno, ti su rezultati protumačeni kao pokazatelj da je neki oblik teorije uma prisutan i prije nego što djeca navršu četiri godine (Baillargeon i sur., 2010; Scott, 2017).

**Podatci iz istraživanja s odraslim osobama.** Prema tradicionalnome gledištu teorije uma odrasle osobe posjeduju teoriju uma, s time da je riječ o kognitivno zahtjevnome sustavu koji je pod svjesnom kontrolom osobe koja ga koristi. Takvim kontroliranim procesima obično se suprotstavljaju automatski procesi koji se događaju i pokreću bez kontrole ispitanika (Shiffrin i Schneider, 1977).

Tradicionalno gledište teorije uma implicira da procesiranje mentalnih stanja drugih osoba ne bi trebalo postojati izvan svjesne kontrole i izvan situacija u kojima su mentalna stanja drugih osoba relevantna za sam zadatak. Međutim, postoje istraživanja koja pokazuju da tuđa mentalna stanja mogu utjecati na prosudbe ispitanika u zadacima u kojima ne postoji poseban razlog za obradu tuđega mentalnog stanja. U takvim zadacima ne postoji razlog zašto bi ispitanici uopće pokrenuli proces obrade tuđih mentalnih stanja. Paradigma koju koriste Kovács i suradnici (2010; predstavljena u prethodnome odlomku) slijedi tu logiku te uključuje prikazanu osobu koja posjeduje uvjerenje koje se ili podudara ili se ne podudara s vjerovanjem djeteta. Kovács i suradnici (2010) također izvješćuju o sličnome istraživanju provedenome s odraslim osobama. Ispitanicima je rečeno da pritisnu gumb kada detektiraju da se, nakon što je pregrada uklonjena, iza pregrade nalazi lopta. Manipulirano je vjerovanjem ispitanika o tome hoće li lopta biti prisutna ili odsutna nakon što se ukloni pregrada. Kriična se usporedba odnosila na dva uvjeta u kojima je ispitanik uvijek očekivao da lopte neće biti, ali se ona uvijek tamo nalazila nakon što je pregrada uklonjena, no uvjeti su se razlikovali u onome što je prikazana osoba očekivala. Rezultati te kritične usporedbe pokazali su da su ispitanici, unatoč tomu što je njihovo vjerovanje bilo isto u obama uvjetima, brže reagirali kada je prikazana osoba vjerovala da će lopta biti prisutna nego kada je osoba, kao i oni sami, vjerovala da lopta neće biti prisutna.

Povezana istraživanja pokazala su da se slična interferencija događa kada ispitanici trebaju prosuditi o onome što vide, a kada je prisutna i druga osoba koja je prikazana u samome zadatku. Jedna je od najčešće korištenih paradigmi u tome području tzv. *zadatak zauzimanja perspektive s točkama* (engl. *dot perspective taking task*; Samson i sur., 2010). Ispitanicima je prikazana slika sobe u kojoj se vide lijevi i desni zid. U sredini sobe nalazi se osoba koja je okrenuta ili prema lijevome ili prema desnom dijelu sobe. Prikazane su i točke koje se nalaze ili na lijevome zidu ili na desnome zidu, ili pak na obama zidovima. U nekim pokušajima prikazana osoba stoga vidi jednak broj točaka kao i ispitanik (tj. sve su točke prikazane na zidu prema kojemu je osoba okrenuta), dok u drugim pokušajima osoba vidi različit broj točaka od ispitanika (tj. točke su prikazane na obama zidovima, a prikazana osoba vidi samo one točke koje se nalaze na zidu prema kojemu je okrenuta). Zadatak je ispitanika procijeniti broj točaka koje vide. Ključno je da njihov zadatak ne uključuje procjenu o tome što prikazana osoba vidi, već samo procjenu onoga što oni sami vide. Unatoč uputi i nerelevantnosti perspektive prikazane osobe ispitanici su brže reagirali kada osoba vidi jednak broj točaka kao i oni nego kada osoba vidi različit broj točaka od njih (Samson i sur., 2010). Samson i suradnici taj učinak nazivaju altercentričnom interferencijom (Samson i sur., 2010).

Uz gore opisane paradigme koje su utemeljene na mjerenju vremena reakcije, postoje i drugi pristupi, kao što je upotreba praćenja pokazivača na ekranu dok ga ispitanik pomiče pomoću računalnoga miša tijekom rješavanja zadatka. Van der Wel i suradnici (2014) mjerili su putanju kretanja pokazivača ispitanika prema ciljnome

predmetu za koji je otkriveno da se nalazi na jednoj od dviju mogućih lokacija. Mjerenjem koliko su putanje prema točnoj lokaciji izravne (mjereno kroz područje ispod krivulje) može se ispitati u kojoj mjeri konkurentna lokacija „povlači” ispitanike prema sebi. Putanje pokazivača ispitanika koji nisu dobili upute da trebaju pratiti vjerovanja osobe prikazane na sceni bile su više „povučene” prema netočnoj lokaciji kada je osoba vjerovala da je predmet na toj lokaciji nego kada je osoba imala jednako vjerovanje kao i ispitanik. Slični efekti interferencije također su opaženi u istraživanjima zauzimanja perspektive. Primjerice, u istraživanju Marshall i suradnika (2018) ispitanici su morali naznačiti koliko predmeta vide, i to pomoću ljestvice koja je imala oznake samo na svom početku i kraju (početak: 0; kraj: 15), što je ispitanicima otežavalo razlikovanje apsolutne pozicije njihova pokazivača na ljestvici. Ispitanicima je bila prikazana scena u kojoj je osoba mogla vidjeti ili sve predmete na sceni ili samo neke od njih (npr. ispitanik je vidio četiri, a prikazana osoba samo tri predmeta). Kada je osoba vidjela manje predmeta od ispitanika, ispitanici su kliknuli bliže početku ljestvice (dakle, označili su manji broj) nego kada je osoba vidjela jednak broj predmeta kao i oni sami. Dakle, na fizičke prosudbe odraslih ljudi može utjecati ono što druga osoba može vidjeti ili u što vjeruje i onda kada su perspektiva, tj. vjerovanje te druge osobe nevažni.

### **Teorijska objašnjenja izvedbe na zadatcima teorije uma**

Postoje tri glavna tipa teorija za objašnjenje različitih empirijskih rezultata. Sva tri tipa teorija dovode u pitanje tradicionalno gledište teorije uma, međutim, razlikuju se u tome na osnovi čega to čine.

#### **Teorije koje su utemeljene na jednome sustavu (engl. *single system theories*)**

Prema zagovornicima teorija koje su utemeljene na jednome sustavu ne postoji temeljna razlika između mehanizma koji koristimo u verbalnome zadatku pogrešnoga vjerovanja i mehanizma koji koristimo u neverbalnome zadatku pogrešnoga vjerovanja. Nalazi koji upućuju na prisutnost automatske teorije uma koja se rano razvija interpretiraju se kao izravan dokaz da je sustav teorije uma prisutan i u djetinjstvu i da može djelovati s ograničenim izvršnim funkcijama. Prema tome teorijskom objašnjenju razlika između nalaza prema kojima je teorija uma kognitivno zahtjevnija i podataka prema kojima je teorija uma automatski proces koji se rano razvija rezultat je razlike u zahtjevima zadataka: loša izvedba na verbalnim zadatcima pogrešnoga vjerovanja dojenčadi (kao i pogreške odraslih) objašnjavaju se dodatnim zahtjevima takvih zadataka u usporedbi s neverbalnim zadatcima. Kod verbalnih zadataka težina zadatka vjerojatno proizlazi iz toga što djeca moraju povezati višestruke procese ili sustave da bi dala točan odgovor i toga što su određene kombinacije teže od drugih (bilo zato što točan odgovor zahtijeva povezivanje više sustava ili zato što su sami sustavi zahtjevniji; v. npr. Carruthers, 2013). Ta teorija također pruža objašnjenje za različite rezultate povezane s teorijom uma kod odraslih

osoba. Dok su za odrasle ispitanike određene kombinacije kognitivnih sustava jednostavne i mogu se brzo izvesti (pa stoga omogućuju automatske odgovore na neke značajke mentalnih stanja), druge je kombinacije teže izvesti te je vjerojatnije da će one rezultirati pogreškama (Carruthers, 2017).

Ideja da zahtjevi zadataka potencijalno utječu na dječju izvedbu na zadacima pogrešnoga vjerovanja nije nova i prethodi neverbalnim zadacima. Naime, mnogi istraživači koji se bave kognitivnim razvojem već dugo uzimaju u obzir da izvedba raznih verzija verbalnoga zadatka pogrešnoga vjerovanja ovisi ne samo o teoriji uma, nego i o drugim sposobnostima (Chandler i sur., 1989; Leslie, 1994). Primjerice, u takvim zadacima djeca moraju odgovoriti na više pitanja ili imati na umu protučinjenične informacije (Bloom i German, 2000; Leslie, 1994; Riggs i sur., 1998). Nadalje, sama priroda većine zadataka pogrešnoga vjerovanja uključuje kontrast između stvarnosti i znanja prikazane osobe – tako da ispravan odgovor zahtijeva od ispitanika da inhibiraju vlastito znanje i stoga uključe procese koji zahtijevaju izvršne funkcije. Zbog toga neki istraživači (poznati kao *boosters*) tvrde da verbalna verzija zadatka nije dobar test za ispitivanje dobi u kojoj se teorija uma razvija te da teorija uma može biti prisutna kod djece mlađe od četiri godine (Chandler i sur., 1989). Ti su argumenti sada dodatno potkrijepljeni podacima koji pokazuju da djeca uspješno rješavaju zadatke pogrešnoga vjerovanja u kojima se mjere neverbalni odgovori i onda kada njihove verbalne verzije ne mogu uspješno riješiti.

### **Teorije koje su utemeljene na nementalističkim procesima**

Drugačija interpretacija empirijskih podataka, a koja također dovodi u pitanje tradicionalno gledište teorije uma, utemeljena je na tvrdnji da izvedba na zadacima u kojima se mjere neverbalni odgovori nema nikakve veze s teorijom uma (Heyes, 2014b, 2015; Ruffman i Perner, 2005). Slično je gledište poznato u području koje se bavi komparativnom kognicijom, gdje se nalazi o teoriji uma kod životinja često odbacuju i interpretiraju kao rezultat nekoga drugoga kognitivnog mehanizma (Heyes, 1998; Penn i Povinelli, 2007). Stoga su sve tri naknadno predstavljene verzije takvih nementalističkih teorija općenito utemeljene na ideji da uspješna izvedba na zadatku nije nužno dokaz složene kognitivne sposobnosti – npr. teorije uma – ako se uspješna izvedba može objasniti jednostavnijom kognitivnom sposobnošću (tzv. Morganov kanon; Morgan, 1894). Taj pristup zahtijeva: i) opis jednostavnijega mehanizma, ii) pružanje dokaza da opisani mehanizam postoji i iii) testiranje je li izvedba ispitanika u skladu s predviđanjima toga jednostavnijega alternativnog objašnjenja.

### **Objašnjenja koja se odnose na značajku drugoga mehanizma**

Prema jednome obliku nementalističkoga objašnjenja značajke osnovnih mehanizama učenja, perceptivnih procesa i sustava pamćenja mogu objasniti



rezultate specifičnih eksperimenata (ili skupova eksperimenata) bez potrebe za impliciranjem bilo kojega oblika atribucije mentalnoga stanja (Heyes, 2014a, 2017). Kao primjer možemo navesti istraživanje Onishi i Baillargeon (2005) u kojemu je korištena paradigma kršenja očekivanja da bi se ispitalo rasuđivanja o pogrešnim vjerovanjima kod dojenčadi. Perner i Ruffman (2005) navode da je dojenčad mogla najprije stvoriti asocijaciju između prikazane osobe, predmeta i kutije u kojoj se predmet nalazi te da je naknadno možda dulje gledala u novi događaj nego u kontrolni događaj jer je u novome događaju stvarala novu asocijaciju između tih triju elemenata, dok u kontrolnome uvjetu nije morala stvarati novu asocijaciju. Dakle, razlika u vremenu gledanja između uvjeta točnoga (kontrolni uvjet) i pogrešnog vjerovanja (eksperimentalni uvjet) može se objasniti bez impliciranja teorije uma.

Heyes (2014a) nudi općenitije objašnjenje toga što se događa u paradigmatama koje se koriste s dojenčadi. Njezino se objašnjenje fokusira na sjećanje dojenčadi o događajima koji su prikazani u zadatku (v. također Heyes, 2017 za blisko objašnjenje izvedbe čimpanza u zadatcima koji mjere anticipativno gledanje) i utemeljeno je na fenomenu retroaktivne interferencije. Kod retroaktivne se interferencije sjećanje na događaj smanjuje kada se neki salijentni događaj dogodi ubrzo nakon toga. U većini se uvjeta pogrešnoga vjerovanja prije nego što prikazana osoba treba naći skriveni predmet dogodi neki salijentni događaj – npr. često se osoba vraća na scenu ili se okreće. Heyes (2014a) tvrdi da će takvi događaji vjerojatno ometati sjećanje ispitanika na stvarnu lokaciju na kojoj se predmet nalazi te će povećati relativnu snagu prethodne lokacije predmeta (one za koju osoba pogrešno vjeruje da sadrži predmet) u usporedbi s uvjetom točnoga vjerovanja gdje se osoba ne vraća na scenu (ili gdje predmet zapravo i nije promijenio lokaciju).

### ***Objašnjenja koja su utemeljena na „čitanju ponašanja” (engl. behaviour-reading)***

Izraz *čitanje ponašanja* prvi su upotrijebili Premack i Woodruff (1978), a otada je u fokusu detaljnih rasprava u području komparativne kognicije (Heyes, 1998; Lurz, 2011; Penn i Povinelli, 2007). Prema objašnjenju koje je utemeljeno na čitanju ponašanja zadatci teorije uma mogu se riješiti pomoću znanja o tome da određeno ponašanje prethodi drugomu ponašanju. Takvo učenje zaobilazi potrebu da ispitanici rasuđuju o mentalnim stanjima da bi predvidjeli sljedeće radnje druge osobe jer je njezina sljedeća radnja povezana s njezinom trenutačnom radnjom i/ili kontekstom u kojemu se ona odvija. Pobornici objašnjenja koje je utemeljeno na čitanju ponašanja ističu da se vidljivo ponašanje uvijek može koristiti za predviđanje ponašanja druge osobe jer se mentalna stanja ne mogu direktno promatrati (Lurz, 2009; Povinelli i Vonk, 2004).

Slična objašnjenja postoje i za izvedbu dojenčadi u neverbalnim zadatcima pogrešnoga vjerovanja (Ruffman, 2014; Ruffman i Perner, 2005). Tako se dječja predviđanja ponašanja druge osobe u takvim zadatcima mogu interpretirati kao rezultat asocijativnih ili statističkih procesa učenja za koje je poznato da omogućuju

brzo učenje kontingentnosti u drugim domenama, npr. prilikom učenja jezika (Saffran i sur., 1996).

### ***Objašnjenja koja su utemeljena na procesima pažnje***

Ljudi često reaguju na socijalne podražaje drugačije nego na nesocijalne podražaje. Položaj očiju i/ili smjer u kojemu gleda druga osoba utječu na to gdje će i odrasli i dojenčad usmjeriti pažnju (Friesen i Kingstone, 1998; Hood i sur., 1998). Taj se učinak pojavljuje i kada je riječ o pamćenju, pri čemu ispitanici bolje pamte podražaje u koje je druga osoba gledala od podražaja u koje ta osoba nije gledala (Gregory i Jackson, 2017). Takvi se učinci koriste za različita nementalistička objašnjenja izvedbe na zadacima teorije uma jer ti zadatci često manipuliraju onime što prikazana osoba vidi ili u što vjeruje tako da se manipulira vizualnim pristupom osobe nekome kritičnom događaju.

Falck i suradnici (2014) pozivaju se na procese pažnje da bi objasnili izvedbu dojenčadi na određenim zadacima u kojima se mjeri vrijeme gledanja. Tvrde da objašnjenje za izvedbu na zadacima poput onoga koji koriste Kovács i suradnici (2010) leži u fenomenu koji nazivaju *interesnom zarazom* (engl. *interest contagion*). Dojenčadi su možda zanimljiviji događaji na koje druga osoba obraća pažnju te takve događaje bolje pamte (takvi se učinci nalaze i kod odraslih osoba, v. Gregory i Jackson, 2017). U istraživanju Kovács i suradnika (2011) vrijeme gledanja dojenčadi bilo je kraće u pokušajima kada se dogodio neočekivani događaj ako je osoba koja je prikazana u sceni očekivala taj događaj. Očekivanjima osobe manipuliralo se tako da je ona ili bila prisutna ili nije bila prisutna kada je lopta ušla u scenu, ili se pak udaljila s nje. Prema Falck i suradnicima (2015) razlika u vremenu gledanja može biti rezultat toga što dojenčad ima snažnije kodiranje lopte koja ponovno ulazi u scenu kada je osoba vidjela događaj jednostavno zato što je osoba bila prisutna pa je stoga dojenčad obratila pažnju na taj kritični događaj.

Slična objašnjenja koja su utemeljena na pažnji prisutna su i u literaturi o teoriji uma kod odraslih osoba. Primjerice, pogled prikazane osobe skrenut će pažnju ispitanika u smjeru u kojemu osoba gleda (Friesen i Kingstone, 1998). Zadatak zauzimanja perspektive s točkama uključuje sličan scenarij kao i paradigme praćenja pogleda: postoji središnji socijalni podražaj koji je usmjeren prema jednoj strani. Stoga se navodi da bi skretanje pažnje na jednu stranu moglo biti objašnjenje rezultata, a ne zauzimanje perspektive prikazane osobe (Heyes, 2014b; Santiesteban i sur., 2014). Stoga utjecaj socijalnih podražaja na pažnju pruža dobro objašnjenje i za izvedbu odraslih osoba i za izvedbu dojenčadi u neverbalnim zadacima teorije uma.

### ***Teorije koje su utemeljene na dvama sustavima (engl. two-systems theories)***

Treći tip teorija koje pokušavaju objasniti rezultate dojenčadi i odraslih osoba, a koji dovodi u pitanje tradicionalno gledište teorije uma, utemeljen je na

pretpostavci da postoje dva sustava teorije uma. Među takvim objašnjenjima ističe se ono Apperlyja i Butterfilla (2009) koji su izgradili model mehanizma *minimalne teorije uma* (engl. *minimal Theory of Mind*) koji bi mogao djelovati paralelno sa sustavom *potpune teorije uma* (engl. *full-blown Theory of Mind*; Apperly i Butterfill, 2009; Butterfill i Apperly, 2013).

U Apperlyjevoj i Butterfillovoj teoriji *potpuna* teorija uma obrađuje mentalna stanja kao integrirane reprezentacije koje sadrže sve aspekte mentalnih stanja te je u skladu s tradicionalnim gledištem o teoriji uma, kognitivno je zahtjevna, spora i vjerojatno se kasno razvija. Nasuprot tomu, autori predlažu da minimalna teorija uma procesira mentalna stanja putem zamjenika koji ovise o određenim fizičkim odnosima između osobe i vanjskoga svijeta (Apperly i Butterfill, 2009; Butterfill i Apperly, 2013). Rasuđivanje o takvim odnosima nije jednako složeno kao rasuđivanje o reprezentaciji svih aspekata mentalnih stanja osobe te stoga omogućuje brze odgovore i razvoj u ranoj dobi. Međutim, cijena smanjenja zahtjeva procesiranja je proizvodnja netočnih predviđanja u određenim situacijama jer zamjenici za mentalna stanja ne uspijevaju obuhvatiti cijelu složenost mentalnih stanja (Butterfill i Apperly, 2013).

U osnovi teorije Apperlyja i Butterfilla nalazi se oblik pripisivanja ciljeva na niskoj razini, tj. sposobnost prepoznavanja da će druga osoba izvesti radnje koje imaju tendenciju dovođenja do određenih događaja u svijetu. Sljedeća razina proširuje tu sposobnost ograničavanjem situacija u kojima su te radnje predviđene na one u kojima osoba ima određeni fizički odnos s objektima. Npr. očekivala bi se radnja usmjerena prema nekome objektu ako je osoba bila na određenoj udaljenosti od objekta i ako je ta osoba bila okrenuta prema objektu, a scena je bila drugačije osvjetljena. Dakle, ta druga razina, nazvana *susretom* (engl. *encountering*), približava atribuciju percepcije osobe bez potrebe da ispitanik razmišlja o stvarnoj perspektivi osobe. Treća razina, koja se naziva *registracijom* (engl. *registration*), omogućuje ispitanicima da prepoznaju kada će osoba vjerojatno izvesti uspješne radnje usmjerene k cilju jer povezuju ispravno izvođenje radnje s vremensko-prostornim odnosima između osobe i objekta. Npr. ispitanik bi trebao očekivati da će osoba ispravno djelovati na trenutačnoj lokaciji objekta samo ako je posljednji *susret* te osobe s objektom bio na njegovoj trenutačnoj lokaciji. Taj treći dio sustava stoga pruža zamjenike za atribuciju vjerovanja jer povezuje radnje osobe s događajima u kojima je osoba vjerojatno imala perceptivni pristup dotičnomu objektu (izračun perceptivnoga pristupa utemeljen je na drugome dijelu minimalnoga sustava tako da osoba koristi *susret* kao zamjenika za perceptivni pristup). Četvrti element sustava podrazumijeva da rezultat triju prethodnih dijelova minimalne teorije uma ima uzročnu ulogu u radnji osobe. Naime, lokacija na kojoj je osoba *registrirala* objekt navest će osobu da traži objekt na toj lokaciji. Stoga bi predviđanja minimalnoga i potpunog sustava teorije uma trebala biti slična kada ispitanik mora predvidjeti radnje osobe u jednostavnome zadatku pogrešne lokacije. Tu bi minimalni sustav procesirao da je osoba registrirala da je ciljni objekt

postavljen na lokaciju A. Ta se registracija zatim koristi kao osnova za predviđanje radnje osobe unatoč tomu što objekt više nije na lokaciji A.

Prije Apperlyjeve i Butterfillove konceptualizacije sustava minimalne teorije uma (Apperly i Butterfill, 2009; Butterfill i Apperly, 2013) niz prijedloga pozivao se na *srednje* sustave ili procese koji su uključeni u predviđanje tuđih radnji i koji omogućuju kompetentnu izvedbu na nekim zadacima teorije uma bez potrebe za potpunom teorijom uma (Gergely i Csibra, 2003; Whiten, 1996, 2013). Međutim, glavna je prednost Apperlyjeve i Butterfillove teorije u tome što je njihov temeljit opis sustava istaknuo specifične slučajeve u kojima bi upotreba sustava minimalne teorije uma proizvela drugačiji odgovor u odnosu na sustav potpune teorije uma. Stoga njihov prijedlog sadrži specifična ograničenja (engl. *signature limits*) koja se mogu empirijski testirati (Heyes, 2015).

Da bismo razumjeli specifična ograničenja koja predlažu Apperly i Butterfill, korisno je sagledati njihove konceptualizacije dvaju sustava atribucije mentalnih stanja. U teoriji minimalnoga sustava posljednja je lokacija na kojoj neka osoba registrira neki objekt kritični čimbenik u predviđanju radnje te osobe prema tome objektu. Stoga je minimalni sustav dovoljan za točno predviđanje radnji na temelju toga što je osoba vidjela (što se često naziva zauzimanjem perspektive na 1. razini; Flavell i sur., 1976). Međutim, takav sustav ne može dovesti do točnih predviđanja kada je vjerovanje (i radnja) neke osobe utemeljeno na tome *kako* njoj neki predmet izgleda (zauzimanje perspektive na 2. razini). Npr. hoće li osoba reprezentirati objekt kao velik ili mali može ovisiti o tome gleda li u objekt kroz staklo koje djeluje kao povećalo, ali bez da osoba to zna. Ispitanik koji posjeduje potpunu teoriju uma prepoznat će da će na radnje te osobe utjecati to gleda li osoba predmet kroz povećalo ili ne jer povećalo mijenja izgled objekta.

Apperly i Butterfill (2009) objašnjavaju to ograničenje pomoću sljedeće situacije: postoje dvije identične lutke, *lutka A* i *lutka B*. Na početku scene lutka A se – bez da prikazana osoba to zna – nalazi iza pregrade. Osoba gleda kako je lutka B smještena iza pregrade. Osoba zatim promatra kako se lutka A najprije uklanja iza pregrade, a potom i u cijelosti sa scene. Ispitanici s potpunom teorijom uma trebali bi shvatiti da će osoba vjerovati da je jedna jedina lutka bila postavljena iza pregrade te da je zatim uklonjena jer lutka A izgleda identično kao lutka B pa će se osoba prema njoj ponašati kao da je lutka B. Posljedično, osoba bi trebala očekivati da se iza pregrade ne nalazi nijedna lutka te je neće pokušati tražiti na tome mjestu. Suprotno tomu, ispitanik koji koristi minimalnu teoriju uma nije u stanju prepoznati da izgled lutke mijenja razmišljanje prikazane osobe i procesirat će da je osoba registrirala da je lutka B postavljena iza pregrade, a da je lutka A uklonjena. Posljedično, posljednje je mjesto na kojemu je osoba registrirala lutku B iza pregrade te bi minimalni čitač misli predvidio da će je osoba na toj lokaciji i tražiti.

## Testiranje teorijskih objašnjenja

Iako ranije navedena teorijska objašnjenja nude predviđanja o tome kako bi ispitanici trebali reagirati kada se koriste različite manipulacije u zadacima teorije uma, većina je predviđanja zajednička različitim teorijama. Stoga je važno detaljno razmotriti kako bi, prema predviđanjima svake od teorija, obrazac rezultata trebao izgledati.

### Manipulacije u zadacima teorije uma i predviđanja obrasca rezultata

#### *Zadatci s nesocijalnim kontrolama*

Nesocijalni objekti ne mogu posjedovati mentalna stanja. Budući da nementalističke teorije ne priznaju da ispitanici stvarno pripisuju mentalna stanja drugim osobama, jedna je posljedica da bi se ispitanici trebali ekvivalentno ponašati prema socijalnim podražajima (prikazane osobe) i nesocijalnim podražajima, tj. jednaki bi ponašajni učinci trebali biti opaženi kada u zadacima osobu zamijenimo nesocijalnim podražajem. Nasuprot tomu, i teorije koje su utemeljene na jednome sustavu i one koje su utemeljene na dvama sustavima predviđaju da su učinci vidljivi kada se koristi osoba, ali ne i kada se koristi nesocijalni podražaj. Međutim, korisnost je nesocijalne kontrole ograničena zbog sljedećih dvaju argumenata. Prvo, slični obrasci rezultata u socijalnim i nesocijalnim uvjetima često se objašnjavaju antropomorfizmom: ispitanici se možda ponašaju prema nesocijalnome podražaju kao prema osobi, tako da nesocijalni kontrolni uvjet i ne tumače kao takav (Furlanetto i sur., 2016; Surian i Franchin, 2020). Drugo, razlike u obrascu rezultata između socijalnih i nesocijalnih uvjeta mogu se objasniti kao rezultat toga što nesocijalni podražaj nije izravno ekvivalentan socijalnomu podražaju – npr. socijalni podražaji obično drugačije utječu na pažnju ispitanika od nesocijalnih podražaja (Heyes, 2014b).

#### *Zadatci zauzimanja perspektive na 2. razini*

Jedno je od ključnih predviđanja teorija koje su utemeljene na dvama sustavima da ispitanici neće moći točno predvidjeti ponašanje druge osobe čije je vjerovanje manipulirano izgledom objekta kada koriste svoj minimalni sustav teorije uma (Apperly i Butterfill, 2009; Low i sur., 2016). Međutim, slična ograničenja predviđaju i nementalističke teorije kao što su objašnjenja čitanja ponašanja jer ispitanici koji *čitaju* ponašanje mogu oblikovati pravila samo na temelju ponašanja i događaja koje su opazili, a ne na temelju toga *kako* je druga osoba kodirala neki objekt (Lurz, 2009, 2011). Ni teorije koje su utemeljene na jednome sustavu vjerojatno ne predviđaju da će ispitanici točno riješiti zadatke zauzimanja perspektive na 2. razini. Naime, zadatak zauzimanja perspektive na 2. razini zahtijeva od ispitanika da se prisjete izgleda objekta u trenutku kada ga je prikazana osoba vidjela i da inhibiraju odgovor na temelju stvarnoga identiteta objekta. Oba su

procesa vjerojatno kognitivno zahtjevna i sva njihova ograničenja – poput toga da nisu u potpunosti razvijeni – mogu dovesti do toga da ispitanici ne uspiju riješiti zadatak. Čini se, dakle, da sve tri vrste teorije mogu predvidjeti da će ispitanici imati poteškoća sa zauzimanjem perspektive na 2. razini u neverbalnim zadacima teorije uma i stoga takvi empirijski rezultati sami po sebi ne mogu napraviti razliku između tih teorijskih objašnjenja. Međutim, samo teorija koja je utemeljena na dvama sustavima jasno tvrdi da postoji čvrsto ograničenje za zauzimanje perspektive na 2. razini, što znači da je svaki rezultat koji ukazuje na uspješno takvo zauzimanje perspektive u neverbalnim zadacima u sukobu s tom teorijom (Low i sur., 2016).

### **Zadaci projekcije iskustva**

Prema nementalističkim teorijama teorije uma, uključujući objašnjenje utemeljeno na čitanju ponašanja, ispitanici ne mogu koristiti vlastito iskustvo kao osnovu za predviđanje tuđega iskustva (Heyes, 2014a; v. Lurz i Krachun, 2019). Međutim, to se odnosi samo na situacije u kojima je iskustvo o kojemu je riječ novo jer u suprotnome ispitanici mogu koristiti ranije naučene odnose između situacijskih čimbenika i ponašanja drugih osoba da bi predvidjeli kako će se druga osoba ponašati (Heyes, 1998). Glavna je metoda koja se predlaže kao test projekcije iskustva tzv. *zadatak s naočalama* (engl. *goggles task*) u kojemu ispitanik stječe iskustvo noseći dvije različite vrste naočala s kojima prethodno nema iskustva. Jedne naočale omogućuju ispitaniku da vidi, a druge mu blokiraju vid, dok se s vanjske strane (tj. sve dok ih ispitanik ne iskusi na svojim očima) naočale razlikuju samo značajkama koje nisu povezane s njihovim učinkom na vid (npr. razlikuju se u boji i oznakama). Ako ispitanik može projicirati vlastito iskustvo na tuđe, trebao bi moći prepoznati da će prikazana osoba vidjeti s jednim naočalama, a neće vidjeti s drugima te da će to utjecati na ponašanje osobe (Heyes, 1998). Međutim, ključno je da iskustvo ispitanika mora biti novo – ispitanici ne smiju moći povezati neprozirne naočale s drugim neprozirnim predmetima jer, ako mogu stvoriti takvu asocijaciju, mogli bi predvidjeti ponašanje osoba na temelju prethodnoga iskustva o tome kako predmeti koji ometaju vidokrug utječu na ponašanje (Lurz i Krachun, 2019).

U slučaju teorija koje su utemeljene na jednome sustavu sustav bi trebao omogućiti projekciju iskustva kada nema ograničenja u njegovu pristupu kognitivnim mehanizmima koji dopuštaju prisjećanje i korištenje iskustva o kojemu je riječ. Predviđanje teorija koje je utemeljeno na dvama sustavima manje je jasno i ovisi o tome može li minimalni sustav ažurirati okolnosti u kojima obrađuje informacije o tome da se druga osoba *susreće* s događajem ili da ga *registrira* na temelju vlastitoga iskustva ispitanika u takvim situacijama. Npr. kada ispitanik iskusi neprozirne naočale, minimalni sustav može ukloniti sve situacije u kojima netko drugi nosi te naočale s popisa situacija u kojima se bilježi *susret*. Nasuprot tomu, ako minimalni sustav nije sposoban za takvo ažuriranje, ispitanik će se nastaviti ponašati prema osobama koje nose takve naočale kao da se *susreću* s predmetima ispred sebe ili da ih *registriraju*.

## Ključni empirijski nalazi

### *Zadaci s nesocijalnim kontrolama*

**Dojenčad.** Korištenjem nesocijalnih kontrola u istraživanjima teorije uma kod dojenčadi dobivaju se nekonzistentni rezultati. Kovács i suradnici (2010) izvješćuju da se učinak interferencije vjerovanja koji su primijetili kod biheviornalne reakcije kršenja očekivanja kod dojenčadi više ne može detektirati kada se osoba koja se koristi u originalnoj, socijalnoj verziji zadatka zamijeni hrpom kutija. Autori su te rezultate interpretirali tako da učinak interferencije vjerovanja ovisi o tome je li riječ o podražaju u zadatku s intencionalnom osobom ili o nesocijalnome podražaju.

Nasuprot tomu, prema Burnside i suradnicima (2020) obrazac je ponašanja dojenčadi u nesocijalnoj verziji zadatka pogrešnoga vjerovanja koji su koristili Onishi i Baillargeon (2005), a koji se temelji na paradigmi kršenja očekivanja, sličan obrascu u originalnome istraživanju. U toj nesocijalnoj varijanti osoba je zamijenjena dizalicom koja podiže ciljni objekt iz jedne od dviju kutija. Međutim, Surian i Franchin (2020) koriste argument antropomorfizma i predlažu da se dojenčad možda ponašala prema dizalici kao prema podražaju s karakteristikama osobe, a ne kao prema nesocijalnome podražaju. Stoga Surian i Franchin koriste zadatak pogrešnoga vjerovanja koji je utemeljen na anticipativnome gledanju u kojemu manipuliraju time doima li se trokut tijekom faze upoznavanja eksperimenta kao *osoba* ili kao nesocijalni podražaj. Na početku eksperimenta prikazuje se scena potjere u kojoj trokut blisko prati ciljni objekt. U *socijalnome* uvjetu trokut se mogao slobodno kretati bilo kuda na ekranu, a pratila ga je ruka, dok je u *nesocijalnome* uvjetu prikazano kako ta ista ruka pomiče trokut. Stoga bi se, za razliku od nesocijalnoga uvjeta, u socijalnome uvjetu to što trokut prati ciljni objekt moglo protumačiti kao voljna radnja. U testnim pokušajima trokut je prikazan kao da ima pogrešno vjerovanje i mogao se slobodno kretati prema jednoj od dviju lokacija (od kojih je jedna bila lokacija za koju je trokut pogrešno vjerovao da sadrži predmet). I kod odraslih i kod 20-mjesečne dojenčadi autori izvješćuju o dobivenim razlikama u obrascima gledanja između uvjeta: ispitanici su češće gledali prema mjestu za koje je trokut pogrešno vjerovao da sadrži predmet u socijalnome uvjetu nego u nesocijalnome uvjetu. Autori upućuju na to da su ti rezultati u sukobu s teorijom o retroaktivnoj interferenciji prema Heyes (2014) jer bi blisko podudarne perceptivne informacije između uvjeta prema tome objašnjenju trebale proizvesti isti obrazac anticipativnoga gledanja. Međutim, moguće je tvrditi da očiti nedostatak intencionalnosti i voljne radnje u nesocijalnome uvjetu utječe na to koliko pažnje ispitanici usmjeravaju prema prikazanim radnjama i događajima te da ta razlika može objasniti trenutačne rezultate (v. Heyes, 2014 za objašnjenje zašto je zadatak s naočalama prikladniji test nementalističkih teorija od upotrebe nesocijalnih kontrola).

**Odrasle osobe.** Santiesteban i suradnici (2014) zamijenili su osobu korištenu u izvornome zadatku zauzimanja perspektive strelicom koja pokazuje u istome smjeru

kao i osoba. Naime, socijalni podražaj, tj. osoba inducira praćenje pogledom na određeno mjesto, a zamjena osobe strelicom omogućuje korištenje nesocijalnoga podražaja sa sličnim učinkom usmjeravanja. Učinak interferencije socijalnoga podražaja tu je reproduciran nesocijalnim podražajem, što ukazuje na to da je riječ o usmjeravanju pogleda, a ne o atribuciji mentalnih stanja. Slični su rezultati dobiveni i kod verzije zadatka zauzimanja perspektive s točkama u kojemu se ne mjeri vrijeme reakcije (Marshall i sur., 2018). Primjerice, kada su ispitanici trebali označiti koliko predmeta sami vide pomoću neoznačene ljestvice, postavili su oznaku niže na ljestvici kada je podražaj „vidio” samo dio točaka nego kada je „vidio” sve točke prikazane na sceni (Marshall i sur., 2018), i u uvjetu u kojemu je podražaj bila osoba, i u uvjetu u kojemu je podražaj bio trokut.

Međutim, korištenje simbola kojima se prenose informacije (kao što su strelice i trokuti koji pokazuju prema drugim podražajima) kao nesocijalnih podražaja naišlo je i na kritike jer njihova komunikacijska svrha može dovesti do antropomorfnih tumačenja (Westra i sur., 2021). Da bi izbjegli taj problem, Westra i suradnici (2021) zamijenili su osobu/strelicu dvama novim podražajima koji usmjeravaju pogled, a koji su jednaki, osim što se razlikuju u boji, s time da je za jedan ispitanicima rečeno da je izvanzemaljac (živ), a za drugi da je mineral (neživ). U kritičnoj verziji zadatka zauzimanja perspektive s točkama u kojemu ispitanici donose prosudbe o onome što vide – i u kojemu bi središnji podražaj trebao biti irelevantan – došlo je do učinka interferencije samo kada je ispitanicima bilo rečeno da je središnji podražaj izvanzemaljac. Jedno je od objašnjenja da je interferencija potaknuta samo prisutnošću osoba s mentalnim stanjima ili objekata koji se lako antropomorfiziraju. Međutim, i u tome je slučaju još uvijek i usmjeravanje moguće objašnjenje jer su prethodna istraživanja pokazala da do usmjeravanja pomoću novih podražaja dolazi samo ako su ispitanici obaviješteni o tome da je podražaj osoba (Ristic i Kingstone, 2005).

Također, postoje rezultati koji ukazuju na to da nesocijalni događaji mogu objasniti rezultate interferencije vjerovanja kod odraslih ispitanika (Kovács i sur., 2010). U radu u kojemu pokušavaju replicirati rezultate Kovács i suradnika (2010) pomoću zadatka koji mjeri vrijeme reakcije Phillips i suradnici (2015) izvješćuju da su u originalnoj verziji zadatka ispitanici morali pritisnuti jednu tipku kada se prikazana osoba ponovno pojavi na sceni da bi se osiguralo da ispitanici prate što se događa na sceni. Pokušaj replikacije bio je uspješan kada su autori koristili jednaku proceduru, tj. kada su ispitanici morali pritisnuti tipku, međutim, kada to nije bilo dio procedure, rezultati nisu pokazivali učinak interferencije kao što je bio slučaj u originalnome istraživanju. Niz naknadnih eksperimenata pokazao je da se obrazac rezultata dobiven u originalnome eksperimentu dobiva u odsutnosti osobe sve dok se nesocijalni podražaj prezentira u vremenskome okviru koji odgovara vremenu povratka osobe u originalnome istraživanju na koji su ispitanici morali odgovoriti pritiskom tipke. Iako nijedno od teorijskih objašnjenja ne predviđa takve rezultate, čini se da se učinak interferencije vjerovanja dobiven kod odraslih ispitanika može



protumačiti nementalističkim objašnjenjem (ali v. El Kaddouri i sur., 2020 za proturječnu perspektivu).

### **Zadaci zauzimanja perspektive na 2. razini**

**Dojenčad.** U nizu eksperimenata Low i suradnici (Low i sur., 2014; Low i Watts, 2013) prilagodili su zadatak pogrešnoga vjerovanja u kojemu se mjeri anticipativno gledanje tako da je vjerovanje prikazane osobe povezano s identitetom objekta. U jednome od tih istraživanja Low i Watts (2013) prikazali su ispitanicima situaciju koja je vrlo slična onoj koju su Apperly i Butterfill upotrijebili za ocrtavanje ograničenja minimalne teorije uma. U fazi upoznavanja s tim zadatkom prikazana je osoba koja sjedi nasuprot ispitaniku i u svakome pokušaju izražava jasnu preferenciju prema plavim predmetima tako da, kada su prezentirani jedan plavi i jedan crveni predmet, uvijek posegne za plavim predmetom. Tijekom testnih pokušaja osoba i dalje sjedi nasuprot ispitaniku. Prikazuje se robot s različito obojenim stranama, s time da je plava strana vidljiva osobi, a crvena ispitaniku. Robot se pomiče dok ne dođe iza kutije koja prikazanoj osobi onemogućuje da ga vidi. U tome ključnom trenutku robot se okreće za 180 stupnjeva i ispitanik tada vidi stranu koja je ranije bila vidljiva osobi te tako zna da robot ima jednu plavu i jednu crvenu stranu. Nakon što robot iziđe iza kutije, pomiče se iza druge kutije. Stoga prikazana osoba sada ima pogrešno vjerovanje da postoje dva robota, jedan crveni i jedan plavi, te da je plavi robot ostao iza jedne kutije, dok se crveni robot pomaknuo iza druge kutije. U skladu s predviđanjima minimalne teorije uma ispitanici nisu značajno dulje gledali prema točnoj lokaciji (tj. prema kutiji za koju osoba vjeruje da se iza nje nalazi plavi robot) negoli je očekivano slučajem. Osim toga, mjesto prema kojemu su ispitanici najprije gledali razlikovalo se od mjesta prema kojemu su ispitanici gledali kada su testirani u originalnome zadatku pogrešnoga vjerovanja pomoću paradigme anticipativnoga gledanja (zadatak zauzimanja perspektive na 1. razini), što upućuje na to da na anticipativno gledanje u scenarijima pogrešnoga vjerovanja utječe ono što je osoba vidjela, ali ne i kako je osoba neki predmet identificirala.

Rezultati dobiveni pomoću paradigme kršenja očekivanja naizgled proturječe onima dobivenima paradigmom anticipativnoga gledanja jer ukazuju na to da dojenčad može prepoznati kada je netko zavaran izgledom predmeta (14-mjesečna dojenčad, Song i Baillargeon, 2008; 18-mjesečna dojenčad, Scott i sur., 2010). Međutim, nejasno je pokazuju li ti rezultati da dojenčad razmišlja tome *kako predmeti izgledaju* nekoj drugoj osobi (Butterfill i Apperly, 2013; Scott i sur., 2015). Glavni argument protiv zauzimanja perspektive na 2. razini je da dojenčad možda razmišlja o *vrstama* dostupnih objekata jer je svjesna da se drugoj osobi uvijek prikazuju dvije vrste objekata i da osoba preferira određenu vrstu (objekt s plavom kosom, objekt koji je trenutačno rastavljen na dva dijela). Rasuđivanje o *vrsti* objekta omogućilo bi dojenčadi da ispravno odgovori bez potrebe da razmišlja o tome kako će izgled objekta utjecati na osobu (Butterfill i Apperly, 2013).

**Odrasle osobe.** Surtees i Apperly (2012) izmijenili su zadatak zauzimanja perspektive s točkama koji su koristili Samson i suradnici (2010) tako da su ispitanici trebali donositi procjene o magnitudama koje su prikazivale znamenke (npr. je li prikazani broj veći od 7). Kada se koriste znamenke slične onima na digitalnome satu, asimetrični brojevi poput 6 i 9 predstavljaju različite magnitude kada se gledaju iz različitih kutova, ali simetrični brojevi poput 8 i 5 predstavljaju istu magnitudu iz svih kutova. Način na koji bi osoba prikazana na sceni percipirala predstavljenu znamenku nije utjecao na izvedbu ispitanika (npr. Surtees i Apperly, 2012). Slični su rezultati pronađeni kod varijanti zadataka u kojima se mjerilo vrijeme reakcije i u kojima se manipuliralo vjerovanjem osobe o prisutnosti predmeta i načinu kako osoba vidi predmet (Edwards i Low, 2019).

Međutim, drugi eksperimenti dobivaju drugačije rezultate, tj. dobivaju rezultate koji ukazuju na to da način kako drugi percipiraju objekt može proizvesti efekt altercentrične interferencije (Elekes i sur., 2016; Surtees i sur., 2016). U obama je eksperimentima korišten zadatak prepoznavanja brojeva sličan onomu koji su koristili Surtees i Apperly (2012), ali je umjesto prikazivanja scene i osobe na ekranu računala eksperiment proveden *uživo*. Vodoravno postavljen ekran prikazivao je relevantne znamenke (koje su bile ili simetrične (0, 8, 5) ili asimetrične (6, 9)). Ispitanici su testirani ili zajedno s drugim ispitanikom koji je stajao na suprotnoj strani monitora ili pak pojedinačno. U obama istraživanjima ispitanici su opet trebali donositi procjene o magnitudama prikazanih znamenki. Učinak u kojemu su ispitanici sporije rasuđivali o magnitudi prikazane znamenke kada je znamenka bila asimetrična nego kada je bila simetrična bio je veći u uvjetu u kojemu su bili testirani zajedno s drugim ispitanikom nego u uvjetu u kojemu su bili testirani sami – tj. došlo je do efekta altercentrične interferencije. U tim dvama istraživanjima dodatno je manipulirano zadatkom koji je dodijeljen partneru, tj. drugomu ispitaniku – zadatak je bio povezan ili s magnitudom znamenke ili s nekom drugom značajkom znamenke (boja prikazane znamenke, Elekes i sur., 2016; je li prikazana znamenka imala točkice na sebi, Surtees i sur., 2016). Tu su se rezultati dvaju istraživanja razlikovali: Elekes i suradnici (2016) nisu dobili statistički značajnu razliku u vremenu reakcije između asimetrične i simetrične znamenke kada partnerov zadatak nije bio povezan s magnitudom prikazane znamenke, dok Surtees i suradnici (2016), koristeći se sličnom manipulacijom, izvješćuju da je dobiven efekt altercentrične interferencije. Nadalje, efekt altercentrične interferencije dobiven je i u pokušajima u kojima partner nije trebao odgovoriti dok su god ispitanici bili svjesni da je partner prethodno sudjelovao u prosudbama o magnitudi (Surtees i sur., 2016).

### ***Zadatci projekcije iskustva***

**Dojenčad.** Varijante zadatka s naočalama korištene su kod dojenčadi zajedno s paradigmom anticipativnoga gledanja. Senju i suradnici (2011) koristili su zadatak pogrešnoga vjerovanja u kojemu su mjerili anticipativno gledanje tako da je jedna grupa ispitanika iskusila povez kroz koji se moglo vidjeti, tj. *proziran povez*, a druga

grupa povez kroz koji se nije moglo vidjeti, tj. *neproziran povez*. Nakon što je iskusila ili prozirni ili neprozirni povez, dojenčad je dobila tipičan zadatak predviđanja pogrešnoga vjerovanja, ali, umjesto da prikazana osoba napusti scenu, osoba je stavila povez na oči i čekala dok se ciljni objekt pomicao. U skladu s posjedovanjem sustava mentalizacije, dojenčad je promijenila svoj obrazac gledanja ovisno o svojemu prošlom iskustvu. Dojenčad koja je iskusila *neprozirni povez* očekivala je da će osoba izvesti radnju na originalnoj lokaciji objekta, dok taj učinak nije dobiven kod dojenčadi koja je iskusila *prozirni povez*. Slični su rezultati dobiveni u istraživanju s trogodišnjacima u kojemu se mjerilo anticipativno gledanje, a prikazana je osoba nosila sunčane naočale (Király i sur., 2018). Međutim, Heyes (2014) tvrdi da su rezultati koje su objavili Senju i suradnici (2011) neuvjerljivi jer povezi za oči nisu stavljeni izravno preko očiju dojenčadi nego samo blizu očiju, što im je omogućilo da prepoznaju sličnosti između poveza za oči i različitih vrsta objekata koje su prethodno iskusili, a što im je pak omogućilo predviđanje radnji osobe na temelju generalizacije ranije naučenih odnosa između objekata koji se nalaze u vidnome polju i prekrivaju druge predmete (u tome slučaju vrste poveza za oči) i ponašanja u takvim situacijama, a ne stvarni utjecaj objekta na sposobnost osobe da vidi kroz povez. Sličan argument vjerojatno vrijedi i za zadatak koji su koristili Király i suradnici (2018).

**Odrasle osobe.** Prilagodbe zadatka s naočalama korištene su i u kombinaciji sa zadatkom zauzimanja perspektive s točkama, ali dobiveni su rezultati nejasni. Furlanetto i suradnici (2016) otkrili su manji učinak altercentrične interferencije kada je osoba prikazana s naočalama koje je ispitanik iskusio kao *neprozirne* nego kada je osoba prikazana s naočalama koje je ispitanik iskusio kao *prozirne*. U istraživanju Conwaya i suradnika (2017) ispitanici su u zadatku zauzimanja perspektive s točkama trebali prosuditi o broju točaka samo iz vlastite perspektive, a postojale su dvije varijante projekcije iskustva. Jedna je varijanta koristila *teleskopske* uređaje kroz koje su ispitanici gledali da bi iskusili promjenu u broju točaka koji su mogli vidjeti. Kroz oba se teleskopa moglo vidjeti, ali se brojem točaka koji se mogao vidjeti manipuliralo pomoću različitih leća teleskopa: jedan je teleskop pokazivao sve točke, a drugi je imao žarišnu duljinu koja je onemogućavala da se vidi ijedna točka. Tijekom samoga zadatka prikazana je osoba gledala kroz jedan ili drugi teleskop. Za razliku od rezultata dobivenih u istraživanju Furlanetta i suradnika (2016), vrijeme reakcije ispitanika nije bilo povezano s brojem točaka koje je osoba mogla vidjeti (Conway i sur., 2017). Nadalje, Conway i suradnici (2017) replicirali su postupak koji su koristili Furlanetto i suradnici (2015) s većim uzorkom ispitanika i nisu potvrdili originalne rezultate, što dodatno upućuje na to da učinak interferencije opažen u zadatku zauzimanja perspektive s točkama nije utemeljen na projekciji iskustva.

## Rasprava

U radu su predstavljene tri glavne teorije koje pokušavaju objasniti obrazac rezultata dobivenih u suvremenim istraživanjima o teoriji uma. Te teorije nude različita objašnjenja očitoga odstupanja u izvedbi na zadacima teorije uma kod odraslih osoba i tijekom ranoga kognitivnog razvoja. Svaka od teorija nudi predviđanja o izvedbi i ograničenjima teorije uma koji se mogu testirati. Međutim, nijedna teorija u potpunosti ne predviđa obrazac rezultata koji stvarno postoji u empirijskoj literaturi, a niz empirijskih rezultata zagovornici različitih teorija drugačije i objašnjavaju.

Teorije koje su utemeljene na jednome sustavu upućuju na to da je isti sustav u podlozi kognitivno učinkovite kompetencije teorije uma koja je prisutna u ranome razvoju i kognitivno zahtjevne kompetencije teorije uma koja se razvija kasnije. Ključno je predviđanje te teorije da će izvedba na zadacima teorije uma biti povezana s pristupom sustava kognitivno zahtjevnim procesima i izvršnim funkcijama. Međutim, ta teorija teško objašnjava zašto se rezultati mnogih paradigmi mogu replicirati kada se socijalni podražaj, tj. osoba prikazana u zadatku, zamijeni nesocijalnim objektom jer bi jedinstveni sustav trebao obrađivati samo osobe, a ne nesocijalne objekte koji nemaju mentalna stanja.

Nementalističke teorije objašnjavaju indikacije ranoga razvoja teorije uma i indikacije kognitivno učinkovite teorije uma pomoću drugih kognitivnih procesa, a ne teorije uma. Stoga su te teorije u skladu s rezultatima koji ukazuju na to da se slični efekti postižu i kada se osobe prikazane u zadacima zamijene nesocijalnim objektima, npr. kada se osoba u zadatku zauzimanja perspektive s točkama zamijeni strelicom. Međutim, ako se rezultati koji ukazuju na projekciju iskustva kod male djece i odraslih pokažu valjanima, to bi predstavljalo izazov za tu teoriju jer sustav koji teorija postulira ne bi smio moći koristiti svoja vlastita mentalna stanja i iskustva da predvidi radnje druge osobe. Nadalje, mješoviti rezultati kod zauzimanja perspektive na 2. razini, posebno u slučaju zadatka s odraslim osobama, nude malu vjerojatnost da jedna generička nementalistička teorija može objasniti i uspjehe i neuspjehe u raznim zadacima.

Prema dominantnoj teoriji koja je utemeljena na dvama sustavima minimalna teorija uma ne može procesirati kako objekti izgledaju drugoj osobi te stoga pruža dobro objašnjenje za neuspjeh u neverbalnim zadacima zauzimanja perspektive na 2. razini. Međutim, ta teorija ne može lako objasniti zašto je izvedba uspješna u sličnim neverbalnim zadacima ako oni sadrže manje promjene, kao što je zamjena slike osobe stvarnom osobom. Nadalje, ostaje nejasno očekuju li teorije koje su utemeljena na dvama sustavima postojanje načina da projekciju iskustva uspješno koristi kognitivno učinkoviti proces, tako da nije moguće prosuditi jesu li ti empirijski rezultati u neverbalnim zadacima koji uključuju projekciju iskustva u skladu s tom teorijom.

Ako želimo prosuditi koja teorija bolje objašnjava trenutačne empirijske rezultate, dolazimo do dvaju dodatnih problema. Prvi je problem što su metode korištene za testiranje određene teorije osporavane, što dovodi i do spora o tome što sami empirijski rezultati znače. U slučaju projekcije iskustva za paradigme korištene u literaturi tvrdi se da nisu usklađene s time kako su te paradigme osmislili zagovornici nementalističkih objašnjenja (npr. v. Heyes, 2014 za opis istraživanja Senju i suradnika, 2011). Slično tomu, za zadatke za koje se čini da testiraju *kako* druga osoba vidi predmet i u kojima je dojenčad pokazala uspješnu izvedbu tvrdi se da zapravo nisu testovi zauzimanja perspektive na 2. razini (npr. v. Butterfill i Apperly, 2013 za opis istraživanja Scott i suradnika, 2010). Nadalje, postoje i tvrdnje da se ispitanici mogu ponašati prema nesocijalnim kontrolama kao prema socijalnim podražajima, čime se negira korist takvoga kontrolnog testa ili uvjeta (Furlanetto i sur., 2016; Surian i Franchin, 2020). Stoga, sve dok se istraživači u tome području ne budu mogli usuglasiti oko toga koja je metodologija potrebna za kritični test svake od teorija, uvijek će postojati prostor za tvrdnju da određena metoda – i naknadni rezultat – nisu zadovoljavajući. Drugi je problem, o kojemu se u ovome radu još nije raspravljalo, kako bi se u literaturi trebalo ponašati prema neuspješnim replikacijama prethodnih rezultata. Primjer takve neuspješne replikacije bio je predstavljen ranije u radu u raspravi o projekciji iskustva u zadatku zauzimanja perspektive s točkama, ali problem je mnogo širi te se mnogi originalni rezultati, posebno u literaturi o razvoju teorije uma, ne mogu replicirati (Kulke i sur., 2018; Powell i sur., 2018; Schuwerk i sur., 2018). Stoga će razvoj najbolje teorije o teoriji uma ovisiti o tome koji se eksperimenti mogu ili ne mogu replicirati, ali i o tome kako će se ovo istraživačko područje odlučiti ponašati prema neuspješnim replikacijama.

### **Zaključak**

U početku se teorija uma smatrala kognitivno zahtjevnom vještinom koja se vjerojatno razvija u dobi od oko četiri godine. Međutim, noviji empirijski nalazi pokazuju da takve sposobnosti možda nisu kognitivno zahtjevne i da djeca mlađa od četiri godine možda posjeduju tu sposobnost. Trenutačno nijedna teorija nije prihvaćena kao sveobuhvatno objašnjenje rezultata o teoriji uma. U ovome su radu identificirane dvije potencijalno povezane prepreke pronalaženju sveobuhvatnoga objašnjenja: dvosmislenost u predviđanjima različitih teorija i neuspjeh u repliciranju prethodnih nalaza. Nužno je da u budućnosti različite teorije ponude jasna i nedvosmislena predviđanja o nalazima koji nisu u skladu s teorijom. Stroža predviđanja trebala bi pomoći istraživačima da bolje razumiju razloge neuspješnih replikacija u literaturi jer će biti očitije odgovaraju li ili ne ti neuspjesi predviđanjima određenih teorija.

## Literatura

- Apperly, I. A. i Butterfill, S. A. (2009). Do humans have two systems to track beliefs and belief-like states? *Psychological Review*, 116, 953–970.  
<https://doi.org/10.1037/a0016923>
- Baillargeon, R., Buttelmann, D. i Southgate, V. (2018). Invited commentary: Interpreting failed replications of early false-belief findings: Methodological and theoretical considerations. *Cognitive Development*, 46, 112–124.  
<https://doi.org/10.1016/j.cogdev.2018.06.001>
- Baillargeon, R., Scott, R. M. i He, Z. (2010). False-belief understanding in infants. *Trends in Cognitive Sciences*, 14(3), 110–118. <https://doi.org/10.1016/j.tics.2009.12.006>
- Bennett, J. (1978). Some remarks about concepts. *Behavioral and Brain Sciences*, 1(4), 557–560. <https://doi.org/10.1017/S0140525X00076573>
- Bloom, P. i German, T. P. (2000). Two reasons to abandon the false belief task as a test of theory of mind. *Cognition*, 77(1), B25–31. [https://doi.org/10.1016/s0010-0277\(00\)00096-2](https://doi.org/10.1016/s0010-0277(00)00096-2)
- Burnside, K., Severdija, V. i Poulin-Dubois, D. (2020). Infants attribute false beliefs to a toy crane. *Developmental Science*, 23(2), e12887. <https://doi.org/10.1111/desc.12887>
- Buttelmann, D., Carpenter, M. i Tomasello, M. (2009). Eighteen-month-old infants show false belief understanding in an active helping paradigm. *Cognition*, 112(2), 337–342. <https://doi.org/10.1016/j.cognition.2009.05.006>
- Butterfill, S. A. i Apperly, I. A. (2013). How to construct a minimal theory of mind. *Mind & Language*, 28(5), 606–637. <https://doi.org/10.1111/mila.12036>
- Carruthers, P. (2013). Mindreading in infancy: Mindreading in infancy. *Mind & Language*, 28(2), 141–172. <https://doi.org/10.1111/mila.12014>
- Carruthers, P. (2017). Mindreading in adults: Evaluating two-systems views. *Synthese*, 194(3), 673–688. <https://doi.org/10.1007/s11229-015-0792-3>
- Chandler, M., Fritz, A. S. i Hala, S. (1989). Small-scale deceit: Deception as a marker of two-, three-, and four-year-olds' early theories of mind. *Child Development*, 60(6), 1263. <https://doi.org/10.2307/1130919>
- Clements, W. A. i Perner, J. (1994). Implicit understanding of belief. *Cognitive Development*, 9, 377–395. [https://doi.org/10.1016/0885-2014\(94\)90012-4](https://doi.org/10.1016/0885-2014(94)90012-4)
- Conway, J. R., Lee, D., Ojaghi, M., Catmur, C. i Bird, G. (2017). Submentalizing or mentalizing in a level 1 perspective-taking task: A cloak and goggles test. *Journal of Experimental Psychology. Human Perception and Performance*, 43(3), 454–465. <https://doi.org/10.1037/xhp0000319>
- Dennett, D. C. (1978). Beliefs about beliefs [P&W, SR&B]. *Behavioral and Brain Sciences*, 1(4), 568–570. <https://doi.org/10.1017/S0140525X00076664>
- Edwards, K., i Low, J. (2019). Level 2 perspective-taking distinguishes automatic and non-automatic belief-tracking. *Cognition*, 193, 104017. <https://doi.org/10.1016/j.cognition.2019.104017>

- El Kaddouri, R., Bardi, L., De Bremaeker, D., Brass, M. i Wiersema, J. R. (2020). Measuring spontaneous mentalizing with a ball detection task: Putting the attention-check hypothesis by Phillips and colleagues (2015) to the test. *Psychological Research*, 84(6), 1749–1757. <https://doi.org/10.1007/s00426-019-01181-7>
- Elekes, F., Varga, M. i Király, I. (2016). Evidence for spontaneous level-2 perspective taking in adults. *Consciousness and Cognition*, 41, 93–103. <https://doi.org/10.1016/j.concog.2016.02.010>
- Epley, N., Keysar, B., Van Boven, L. i Gilovich, T. (2004). Perspective taking as egocentric anchoring and adjustment. *Journal of Personality and Social Psychology*, 87(3), 327–339. <https://doi.org/10.1037/0022-3514.87.3.327>
- Falck, A., Brinck, I. i Lindgren, M. (2014). Interest contagion in violation-of-expectation-based false-belief tasks. *Frontiers in Psychology*, 5. <https://www.frontiersin.org/articles/10.3389/fpsyg.2014.00023>
- Flavell, J. H. (1977). The development of knowledge about visual perception. *Nebraska Symposium on Motivation*, 25, 43–76.
- Friesen, C. K. i Kingstone, A. (1998). The eyes have it! Reflexive orienting is triggered by nonpredictive gaze. *Psychonomic Bulletin & Review*, 5(3), 490–495. <https://doi.org/10.3758/BF03208827>
- Frith, C. i Frith, U. (2005). Theory of mind. *Current Biology*, 15(17), R644–R645. <https://doi.org/10.1016/j.cub.2005.08.041>
- Furlanetto, T., Becchio, C., Samson, D. i Apperly, I. (2016). Altercentric interference in level 1 visual perspective taking reflects the ascription of mental states, not submentalizing. *Journal of Experimental Psychology: Human Perception and Performance*, 42(2), 158–163. <https://doi.org/10.1037/xhp0000138>
- Gergely, G. i Csibra, G. (2003). Teleological reasoning in infancy: The naïve theory of rational action. *Trends in Cognitive Sciences*, 7(7), 287–292. [https://doi.org/10.1016/S1364-6613\(03\)00128-1](https://doi.org/10.1016/S1364-6613(03)00128-1)
- Gilovich, T., Kruger, J. i Savitsky, K. (1999). Everyday egocentrism and everyday interpersonal problems. U: R. M. Kowalski i M. R. Leary (Ur.), *The social psychology of emotional and behavioral problems: Interfaces of social and clinical psychology* (str. 69–95). American Psychological Association. <https://doi.org/10.1037/10320-003>
- Gregory, S. E. A. i Jackson, M. C. (2017). Joint attention enhances visual working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 43, 237–249. <https://doi.org/10.1037/xlm0000294>
- Harman, G. (1978). Studying the chimpanzee's theory of mind. *Behavioral and Brain Sciences*, 1(4), 576–577. <https://doi.org/10.1017/S0140525X00076743>
- Heyes, C. (1998). Theory of mind in nonhuman primates. *Behavioral and Brain Sciences*, 21(1), 101–114. <https://doi.org/10.1017/s0140525x98000703>
- Heyes, C. (2014a). False belief in infancy: A fresh look. *Developmental Science*, 17(5), 647–659. <https://doi.org/10.1111/desc.12148>

- Heyes, C. (2014b). Submentalizing: I am not really reading your mind. *Perspectives on Psychological Science*, 9(2), 131–143. <https://doi.org/10.1177/1745691613518076>
- Heyes, C. (2015). Animal mindreading: What's the problem? *Psychonomic Bulletin & Review*, 22(2), 313–327. <https://doi.org/10.3758/s13423-014-0704-4>
- Heyes, C. (2017). Apes submentalise. *Trends in Cognitive Sciences*, 21(1), 1–2. <https://doi.org/10.1016/j.tics.2016.11.006>
- Hogrefe, G.-J., Wimmer, H. i Perner, J. (1986). Ignorance versus false belief: A developmental lag in attribution of epistemic states. *Child Development*, 57(3), 567–582. <https://doi.org/10.2307/1130337>
- Hood, B. M., Willen, J. D. i Driver, J. (1998). Adult's eyes trigger shifts of visual attention in human infants. *Psychological Science*, 9(2), 131–134. <https://doi.org/10.1111/1467-9280.00024>
- Keysar, B., Barr, D. J., Balin, J. A. i Brauner, J. S. (2000). Taking perspective in conversation: The role of mutual knowledge in comprehension. *Psychological Science*, 11(1), 32–38. <https://doi.org/10.1111/1467-9280.00211>
- Keysar, B., Barr, D. J. i Horton, W. S. (1998). The egocentric basis of language use: Insights from a processing approach. *Current Directions in Psychological Science*, 7(2), 46–49. <https://doi.org/10.1111/1467-8721.ep13175613>
- Keysar, B., Lin, S. i Barr, D. J. (2003). Limits on theory of mind use in adults. *Cognition*, 89(1), 25–41. [https://doi.org/10.1016/S0010-0277\(03\)00064-7](https://doi.org/10.1016/S0010-0277(03)00064-7)
- Király, I., Oláh, K., Csibra, G. i Kovács, Á. M. (2018). Retrospective attribution of false beliefs in 3-year-old children. *Proceedings of the National Academy of Sciences*, 115(45), 11477–11482. <https://doi.org/10.1073/pnas.1803505115>
- Kovács, Á. M., Téglás, E. i Endress, A. D. (2010). The social sense: Susceptibility to others' beliefs in human infants and adults. *Science*, 330(6012), 1830–1834. <https://doi.org/10.1126/science.1190792>
- Kulke, L., Duhn, B. i Schneider, D. (2018). Is implicit theory of mind a real and robust phenomenon? Results from a systematic replication study. *Psychological Science*, 29, 095679761774709. <https://doi.org/10.1177/0956797617747090>
- Leslie, A. M. (1994). Pretending and believing: Issues in the theory of ToMM. *Cognition*, 50(1–3), 211–238. [https://doi.org/10.1016/0010-0277\(94\)90029-9](https://doi.org/10.1016/0010-0277(94)90029-9)
- Lin, S., Keysar, B. i Epley, N. (2010). Reflexively mindblind: Using theory of mind to interpret behavior requires effortful attention. *Journal of Experimental Social Psychology*, 46(3), 551–556. <https://doi.org/10.1016/j.jesp.2009.12.019>
- Low, J., Apperly, I. A., Butterfill, S. A. i Rakoczy, H. (2016). Cognitive architecture of belief reasoning in children and adults: A primer on the two-systems account. *Child Development Perspectives*, 10(3), 184–189. <https://doi.org/10.1111/cdep.12183>
- Low, J., Drummond, W., Walmsley, A. i Wang, B. (2014). Representing How rabbits quack and competitors act: Limits on preschoolers' efficient ability to track perspective. *Child Development*, 85(4), 1519–1534. <https://doi.org/10.1111/cdev.12224>



- Low, J. i Watts, J. (2013). Attributing false beliefs about object identity reveals a signature blind spot in humans' efficient mind-reading system. *Psychological Science*, 24(3), 305–311. <https://doi.org/10.1177/0956797612451469>
- Lurz, R. (2009). If chimpanzees are mindreaders, could behavioral science tell? Toward a solution of the logical problem. *Philosophical Psychology*, 22(3), 305–328. <https://doi.org/10.1080/09515080902970673>
- Lurz, R. (2011). *Mindreading animals: The debate over what animals know about other minds*. MIT Press.
- Lurz, R. i Krachun, C. (2019). Experience-projection methods in theory-of-mind research: Their limits and strengths. *Current Directions in Psychological Science*, 28(5), 456–462. <https://doi.org/10.1177/0963721419850156>
- Marshall, J., Gollwitzer, A. i Santos, L. R. (2018). Does altercentric interference rely on mentalizing?: Results from two level-1 perspective-taking tasks. *Plos One*, 13. <https://doi.org/10.1371/journal.pone.0194101>, e0194101.
- Morgan, C. L. (1894). The scope of psycho-physiology. *Nature*, 49(1274), Article 1274. <https://doi.org/10.1038/049504a0>
- Nickerson, R. (1999). How we know - and sometimes misjudge - what others know: Imputing one's own knowledge to others. *Psychological Bulletin*, 125, 737–759. <https://doi.org/10.1037//0033-2909.125.6.737>
- Onishi, K. H. i Baillargeon, R. (2005). Do 15-month-old infants understand false beliefs? *Science*, 308(5719), 255–258. <https://doi.org/10.1126/science.1107621>
- Penn, D. C. i Povinelli, D. J. (2007). On the lack of evidence that non-human animals possess anything remotely resembling a 'theory of mind.' *Philosophical Transactions of the Royal Society B: Biological Sciences*, 362(1480), 731–744. <https://doi.org/10.1098/rstb.2006.2023>
- Phillips, J., Ong, D. C., Surtees, A. D. R., Xin, Y., Williams, S., Saxe, R. i Frank, M. C. (2015). A Second look at automatic theory of mind: Reconsidering Kovács, Téglás, and Endress (2010). *Psychological Science*, 26(9), 1353–1367. <https://doi.org/10.1177/0956797614558717>
- Povinelli, D. J. i Vonk, J. (2004). We don't need a microscope to explore the chimpanzee's mind. *Mind & Language*, 19, 1–28. <https://doi.org/10.1111/j.1468-0017.2004.00244.x>
- Powell, L. J., Hobbs, K., Bardis, A., Carey, S. i Saxe, R. (2018). Replications of implicit theory of mind tasks with varying representational demands. *Cognitive Development*, 46, 40–50. <https://doi.org/10.1016/j.cogdev.2017.10.004>
- Premack, D. i Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences*, 1(4), 515–526. <https://doi.org/10.1017/S0140525X00076512>
- Pylyshyn, Z. W. (1978). When is attribution of beliefs justified? [P&W]. *Behavioral and Brain Sciences*, 1(4), 592–593. <https://doi.org/10.1017/s0140525x00076895>

- Riggs, K. J., Peterson, D. M., Robinson, E. J. i Mitchell, P. (1998). Are errors in false belief tasks symptomatic of a broader difficulty with counterfactuality? *Cognitive Development*, 13(1), 73–90. [https://doi.org/10.1016/S0885-2014\(98\)90021-1](https://doi.org/10.1016/S0885-2014(98)90021-1)
- Ristic, J. i Kingstone, A. (2005). Taking control of reflexive social attention. *Cognition*, 94(3), B55–65. <https://doi.org/10.1016/j.cognition.2004.04.005>
- Royzman, E. B., Cassidy, K. W. i Baron, J. (2003). “I know, you know”: Epistemic egocentrism in children and adults. *Review of General Psychology*, 7(1), 38–65. <https://doi.org/10.1037/1089-2680.7.1.38>
- Rubio-Fernández, P. i Geurts, B. (2013). How to pass the false-belief task before your fourth birthday. *Psychological Science*, 24(1), 27–33. <https://doi.org/10.1177/0956797612447819>
- Ruffman, T. (2014). To belief or not belief: Children’s theory of mind. *Developmental Review*, 34(3), 265–293. <https://doi.org/10.1016/j.dr.2014.04.001>
- Ruffman, T. i Perner, J. (2005). Do infants really understand false belief? *Trends in Cognitive Sciences*, 9(10), 462–463. <https://doi.org/10.1016/j.tics.2005.08.001>
- Saffran, J. R., Aslin, R. N. i Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, 274(5294), 1926–1928. <https://doi.org/10.1126/science.274.5294.1926>
- Samson, D., Apperly, I. A., Braithwaite, J. J., Andrews, B. J. i Bodley Scott, S. E. (2010). Seeing it their way: Evidence for rapid and involuntary computation of what other people see. *Journal of Experimental Psychology. Human Perception and Performance*, 36(5), 1255–1266. <https://doi.org/10.1037/a0018729>
- Santesteban, I., Catmur, C., Hopkins, S. C., Bird, G. i Heyes, C. (2014). Avatars and arrows: Implicit mentalizing or domain-general processing? *Journal of Experimental Psychology. Human Perception and Performance*, 40(3), 929–937. <https://doi.org/10.1037/a0035175>
- Saxe, R. i Baron-Cohen, S. (2006). Editorial: The neuroscience of theory of mind. *Social Neuroscience*, 1(3–4), 1–9. <https://doi.org/10.1080/17470910601117463>
- Schneider, D., Slaughter, V. P. i Dux, P. E. (2017). Current evidence for automatic theory of mind processing in adults. *Cognition*, 162, 27–31. <https://doi.org/10.1016/j.cognition.2017.01.018>
- Schuwerk, T., Prieswasser, B., Sodian, B. i Perner, J. (2018). The robustness and generalizability of findings on spontaneous false belief sensitivity: A replication attempt. *Royal Society Open Science*, 5(5), 172273. <https://doi.org/10.1098/rsos.172273>
- Scott, R. M. (2017). The Developmental origins of false-belief understanding. *Current Directions in Psychological Science*, 26(1), 68–74. <https://doi.org/10.1177/0963721416673174>
- Scott, R. M. i Baillargeon, R. (2017). Early false-belief understanding. *Trends in Cognitive Sciences*, 21(4), 237–249. <https://doi.org/10.1016/j.tics.2017.01.012>

- Scott, R. M., Baillargeon, R., Song, H. i Leslie, A. M. (2010). Attributing false beliefs about non-obvious properties at 18 months. *Cognitive Psychology*, 61(4), 366–395. <https://doi.org/10.1016/j.cogpsych.2010.09.001>
- Scott, R. M., Richman, J. C. i Baillargeon, R. (2015). Infants understand deceptive intentions to implant false beliefs about identity: New evidence for early mentalistic reasoning. *Cognitive Psychology*, 82, 32–56. <https://doi.org/10.1016/j.cogpsych.2015.08.003>
- Senju, A., Southgate, V., Snape, C., Leonard, M. i Csibra, G. (2011). Do 18-month-olds really attribute mental states to others? A critical test. *Psychological Science*, 22(7), 878–880. <https://doi.org/10.1177/0956797611411584>
- Setoh, P., Scott, R. M. i Baillargeon, R. (2016). Two-and-a-half-year-olds succeed at a traditional false-belief task with reduced processing demands. *Proceedings of the National Academy of Sciences of the United States of America*, 113(47), 13360–13365. <https://doi.org/10.1073/pnas.1609203113>
- Shiffrin, R. M. i Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending and a general theory. *Psychological Review*, 84(2), 127–190. <https://doi.org/10.1037/0033-295X.84.2.127>
- Song, H. i Baillargeon, R. (2008). Infants' reasoning about others' false perceptions. *Developmental Psychology*, 44, 1789–1795. <https://doi.org/10.1037/a0013774>
- Southgate, V., Senju, A. i Csibra, G. (2007). Action anticipation through attribution of false belief by 2-year-olds. *Psychological Science*, 18(7), 587–592. <https://doi.org/10.1111/j.1467-9280.2007.01944.x>
- Surian, L. i Franchin, L. (2020). On the domain specificity of the mechanisms underpinning spontaneous anticipatory looks in false-belief tasks. *Developmental Science*, 23(6), e12955. <https://doi.org/10.1111/desc.12955>
- Surtees, A. D. R. i Apperly, I. A. (2012). Egocentrism and automatic perspective taking in children and adults. *Child Development*, 83(2), 452–460. <https://doi.org/10.1111/j.1467-8624.2011.01730.x>
- Surtees, A., Apperly, I. i Samson, D. (2016). I've got your number: Spontaneous perspective-taking in an interactive task. *Cognition*, 150, 43–52. <https://doi.org/10.1016/j.cognition.2016.01.014>
- van der Wel, R. P. R. D., Sebanz, N. i Knoblich, G. (2014). Do people automatically track others' beliefs? Evidence from a continuous measure. *Cognition*, 130(1), 128–133. <https://doi.org/10.1016/j.cognition.2013.10.004>
- Wellman, H. M., Cross, D. i Watson, J. (2001). Meta-analysis of theory-of-mind development: The truth about false belief. *Child Development*, 72(3), 655–684. <https://doi.org/10.1111/1467-8624.00304>
- Wellman, H. M. i Liu, D. (2004). Scaling of theory-of-mind tasks. *Child Development*, 75(2), 523–541. <https://doi.org/10.1111/j.1467-8624.2004.00691.x>

- Westra, E., Terrizzi, B. F., van Baal, S. T., Beier, J. S. i Michael, J. (2021). Beyond avatars and arrows: Testing the mentalising and submentalising hypotheses with a novel entity paradigm. *Quarterly Journal of Experimental Psychology*, 74(10), 1709–1723. <https://doi.org/10.1177/17470218211007388>
- Whiten, A. (1996). When does smart behaviour-reading become mind-reading? U: P. Carruthers i P. K. Smith (Ur.), *Theories of Theories of Mind* (str. 277–292). Cambridge University Press.
- Whiten, A. (2013). Humans are not alone in computing how others see the world. *Animal Behaviour*, 86(2), 213–221. <https://doi.org/10.1016/j.anbehav.2013.04.021>
- Wimmer, H. i Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition*, 13, 103–128. [https://doi.org/10.1016/0010-0277\(83\)90004-5](https://doi.org/10.1016/0010-0277(83)90004-5)

## **Theory of Mind: A Review of Recent Empirical Results and Main Theoretical Explanations**

### **Abstract**

Theory of Mind is the ability to attribute mental states to others. Until around 20 years ago most evidence pointed to Theory of Mind being a cognitively demanding skill that likely develops at around 4 years of age. However, there is a growing body of literature based on experiments that do not rely on verbal measures that suggests that Theory of Mind-like abilities may occur outside of cognitive control and in infants as young as 9 months. This review discusses the three main types of account that have been proposed as explanations of these new results. Furthermore, it highlights the evidence supporting and contrasting each type of account. There is currently no single account that provides an uncontested explanation of all current data, however, one of the reasons for this is that there is a degree of ambiguity in the predictions and interpretations of each of the accounts making it challenging for any set of experiments to refute an account. Consequently, the future of research on Theory of Mind appears to rely on these accounts producing less ambiguous predictions that cannot be insulated from refutation.

*Keywords:* theory of mind, social cognition, perspective taking

Primljeno: 26. 5. 2023.

