

Nadgledanje razumijevanja narativnoga i ekspozitornoga teksta: Odnos između apsolutne i relativne točnosti metakognitivnih procjena

Igor Bajšanski

Odsjek za psihologiju, Filozofski fakultet Sveučilišta u Rijeci

Sažetak

Točnost se nadgledanja razumijevanja može odrediti na dva načina: relativna se točnost odnosi na korelaciju između procjena i izvedbe na testu razumijevanja, dok se apsolutna točnost odnosi na razliku između procjena i izvedbe. Cilj je ovoga rada bio ispitati odnose između mjera relativne i apsolutne točnosti kod različitih zadataka nadgledanja i mjera razumijevanja teksta. Studenti psihologije (N = 69) su sudjelovali u istraživanju. Čitali su 6 ekspozitornih i 6 narativnih tekstova. Nakon čitanja su procijenili izvedbu na specifičnim vrstama zadataka za ispitivanje razumijevanja. Razumijevanje je svakoga teksta ispitano s 8 zadataka višestrukoga izbora i 4 zadatka otvorenoga tipa. Nakon rješavanja zadataka su procijenili izvedbu. Gama-korelacija je korištena kao pokazatelj relativne točnosti. Indeks pristranosti (razlika između procjene i izvedbe) je korišten kao pokazatelj apsolutne točnosti. Analiza korelacija je pokazala različite obrasce povezanosti za relativnu i apsolutnu točnost. Korelacije su između mjera relativne točnosti u različitim uvjetima (narativni i ekspozitorni tekstovi, pitanja višestrukoga izbora i otvorenoga tipa, nakon čitanja i nakon rješavanja) bile znatno niže od korelacija između mjera apsolutne točnosti. Relativna točnost nije bila povezana s razumijevanjem, dok je apsolutna točnost bila povezana: lošija je izvedba na zadacima razumijevanja bila povezana s pretjeranom samouvjerenošću. Rezultati pokazuju da se relativna i apsolutna točnost razlikuju u stabilnosti, kao i u obrascima povezanosti s razumijevanjem i da uključuju različite metakognitivne procese.

Ključne riječi: metakognicija, metarazumijevanje, nadgledanje razumijevanja, točnost procjena

Uvod

Termin metarazumijevanje označava metakogniciju vezanu za razumijevanje pri čitanju. Metarazumijevanje uključuje dvije komponente: metakognitivno znanje

✉ Igor Bajšanski, Odsjek za psihologiju, Filozofski fakultet Sveučilišta u Rijeci, Slavka Krautzeka bb, 51000 Rijeka. E-pošta: *sibajsan@ffri.hr*

Prikazani su rezultati proizašli iz znanstvenog projekta (Kognitivni, metakognitivni i motivacijski čimbenici čitanja i učenja - 009-1301676-2381) provedenog uz potporu Ministarstva znanosti, obrazovanja i športa Republike Hrvatske.

o čitanju i regulaciju čitanja. Metakognitivno je znanje o čitanju relativno stabilan aspekt metakognicije koji se odnosi na čitateljeva uvjerenja o procesima čitanja, a regulacija dalje uključuje nadgledanje i kontrolu čitanja. Nadgledanje se odnosi na provjeravanje stupnja razumijevanja i uočavanje poteškoća u razumijevanju, dok se kontrola odnosi na primjenu strategija koje mogu poboljšati razumijevanje i omogućiti ostvarivanje postavljenih ciljeva čitanja. Nadgledanje se razumijevanja teksta ispituje pomoću dvije glavne grupe postupaka. Prvo, kod istraživanja se s otkrivanjem pogrešaka od sudionika traži da otkrivaju određene vrste pogrešaka ili nekonzistentnosti u tekstu koji čitaju (Epstein, Glenberg i Bradeley, 1984; Glenberg, Wilkinson i Epstein, 1982). Drugo, metoda se kalibracije razumijevanja odnosi na točnost procjena o razumijevanju pročitanaog teksta, a ta je metoda korištena u ovome radu.

Metoda kalibracije je preuzeta iz psihologije prosuđivanja, u okviru koje je istraživana kalibracija subjektivnih vjerojatnosti na zadacima općega znanja. Kod ovih se istraživanja od sudionika traži da procjene koliko su sigurni da su točno odgovorili na pitanja općega znanja i njihove se procjene uspoređuju s izvedbom.

Osnovni nacrt istraživanja kod metode kalibracije razumijevanja izgleda ovako: sudionici čitaju više tekstova, za svaki tekst procjenjuju koliko su ga razumjeli i zatim rješavaju zadatke kojima se ispituje razumijevanje teksta. Kod ove metode tekstovi ne uključuju nekonzistentnosti, kao kod metode otkrivanja pogrešaka. U seriji su eksperimenata Glenberg i Epstein (1985; Glenberg, Sanocki, Epstein i Morris, 1987) ispitivali nadgledanje razumijevanja pomoću ove metode, a ovakva je osnovna shema primijenjena kasnije u velikom broju istraživanja (npr. Dunlosky, Rawson i Middleton, 2005; Lin, Moore i Zabrocky, 2001; Maki i Serra, 1992a; Maki i Serra, 1992b; Weaver, 1990; Weaver i Bryant, 1995). U istraživanjima u kojima je korištena ova metoda manipulirane su različite varijable kao što su: redoslijed čitanja tekstova, davanja procjena i odgovaranja na pitanja razumijevanja (npr. moguće je čitati prvi tekst, procijeniti, odgovarati na pitanja i zatim prijeći na sljedeći tekst ili je moguće pročitati sve tekstove, procijeniti ih i zatim odgovarati na pitanja), vrste procjena (npr. procjene razumijevanja teksta ili procjene izvedbe na zadacima razumijevanja), obilježja teksta (npr. narativni ili ekspozitorni tekst), obilježja zadataka za ispitivanja razumijevanja (zadaci zaključivanja, zadaci prepoznavanja ili dosjećanja informacija iz teksta, tematski ili specifični zadaci), broj zadataka za ispitivanje razumijevanja, a ispitivane su i osobine sudionika, kao što je prethodno znanje o temi i motivacija.

Temeljni je istraživački problem u ovim istraživanjima odrediti kakav je odnos između procjena i izvedbe na testu razumijevanja. Ovaj je odnos određen kao točnost procjena: ako se procjene podudaraju s izvedbom, te su procjene točne, a ako se ne podudaraju, one su netočne. Najčešći je način prikupljanja procjena kod metode kalibracije razumijevanja ispunjavanjem M-mjesne skale Likertova tipa koja se odnosi na procjenu o razumijevanju i gdje je $M \geq 2$, što je uobičajeno i u istraživanjima metamemorije. Nakon prikupljenih je podataka o procjeni i izvedbi

na testu razumijevanja potrebno odrediti točnost tih procjena. Važno je naglasiti da postoje dva osnovna tipa točnosti: apsolutna i relativna točnost, a to vrijedi i za istraživanja metamemorije.

Kod apsolutne se točnosti ispituju razlike između veličine procjena i izvedbe na testu razumijevanja, odnosno, proporcije točnih odgovora. Kao primjer istraživanja u kojima je ispitivana apsolutna točnost procjena mogu se navesti istraživanja procjena uvjerenja u točnost odgovora na pitanjima općega znanja (npr. Fischhoff, Slovic i Lichtenstein, 1977). U tim se istraživanjima sudionicima prezentiraju tvrdnje koje se odnose na opće znanje i oni za svaku tvrdnju procjenjuju je li točna. Nakon toga na skali od 50% (pogađanje) do 100% (potpuna sigurnost) procjenjuju koliko su sigurni da su točno odgovorili. Visok je stupanj apsolutne točnosti definiran kod velikog broja takvih procjena kao slaganje između određene procjene (npr. 100% ili 80%) i proporcije točnih odgovora kod zadataka koji su tako procijenjeni. Ako je u podskupu čestica koje su procijenjene sa 100% izvedba na testu pamćenja 100%, u podskupu čestica koje su procijenjene s 80% izvedba na testu pamćenja 80% itd., riječ je potpunoj apsolutnoj točnosti. Ovakvi se podaci često prikazuju u obliku krivulje kalibracije kod koje se na apscisi prikazuju pojedine vrijednosti procjene, a na ordinati proporcija točnih odgovora. Osim toga, kao mjera apsolutne točnosti može se izračunati i indeks pristranosti koji se određuje kao razlika između procjene i proporcije točnih odgovora. Pozitivne vrijednosti toga indeksa upućuju na precjenjivanje vlastite izvedbe, a negativne vrijednosti na podcjenjivanje vlastite izvedbe (Schraw, Dunkle, Bendixen i Roedel, 1995).

U istraživanjima je metamemorije i metarazumijevanja veći naglasak stavljen na relativnu, a ne na apsolutnu točnost procjena. U tipičnom se eksperimentu iz područja metamemorije procjenjuje u kojoj mjeri veličina procjena o učenju predviđa izvedbu na testu pamćenja s obzirom na različite eksperimentalne uvjete. U području se metarazumijevanja procjenjuje u kojoj mjeri procjena o razumijevanju predviđa izvedbu na testu razumijevanja u različitim eksperimentalnim uvjetima. Za svakoga se sudionika uspoređuju njegove procjene s izvedbom na testu pamćenja ili razumijevanja izračunavanjem intraindividualne korelacije između procjena i rezultata na testu pamćenja ili razumijevanja. U istraživanjima je metamemorije relativna točnost definirana s obzirom na usporedbu između pojedinih čestica: ako je procjena viša za česticu i u odnosu na česticu j , izvedba bi na testu pamćenja trebala biti viša za česticu i u odnosu na česticu j . Valja naglasiti da visoka relativna točnost ne mora biti praćena i visokom apsolutnom točnošću (Nelson i Narens, 1990). Na primjer, moguće je da je za svaku česticu koje se sudionik uspješno dosjetio procjena u fazi usvajanja bila viša od procjena za svaku česticu koje se sudionik nije dosjetio, što je pokazatelj visoke relativne točnosti procjene, a da je ujedno na svim procjenama znatno podcijenjena ukupna izvedba na testu pamćenja. Veći naglasak na relativnoj točnosti proizlazi iz toga što se sposobnost diskriminiranja između bolje i lošije naučenih čestica, kao i

bolje i lošije savladanih tekstova smatra fundamentalnim obilježjem metakognitivnoga nadgledanja (Nelson i Narens, 1990).

Iako je relativnu točnost metamemorijskih procjena, kao i procjena o razumijevanju moguće računati na više različitih načina (Nelson, 1984), u većini je istraživanja kao mjera te povezanosti korišten Goodman-Kruskallov gama-koeficijent korelacije.

Dva su važna teorijska pitanja koja se odnose na specifičnosti relativne i apsolutne točnosti nadgledanja. Prvo, problem je kakav je odnos između različitih mjera točnosti nadgledanja u različitim situacijama nadgledanja. Drugo, problem je jesu li relativna i apsolutna točnost nadgledanja razumijevanja povezane s razumijevanjem. Ta su dva problema ispitana u ovome radu.

Što se tiče prvoga problema, uobičajena je pretpostavka, često i implicitna, da ljudi posjeduju općenitu sposobnost nadgledanja kognitivnoga procesiranja koja će se manifestirati kod različitih kognitivnih zadataka. Ako je to tako, ljudi bi s boljom općom sposobnošću nadgledanja trebali pokazati i viši stupanj točnosti nadgledanja razumijevanja. Međutim, malo je istraživanja u kojima je eksplicitno ispitivano postojanje takve općenite sposobnosti nadgledanja.

Schraw i sur. (1995) su ispitivali nadgledanje izvedbe u različitim domenama kako bi odgovorili je li sposobnost nadgledanja specifična s obzirom na pojedine domene ili je ona općenita metakognitivna sposobnost. Njihovi su sudionici rješavali zadatke iz osam različitih domena (razumijevanje pri čitanju, udaljenosti između američkih gradova, energetska vrijednost hrane, američki predsjednici, brzina kretanja životinja, prostorne procjene, opće znanje, matematički zadaci riječima) i procjenjivali izvedbu na tim zadacima. Ovi zadaci uključuju različite vrste kognitivnih procesa i razlikuju se u vrsti i količini prethodnoga znanja potrebnoga za njihovo rješavanje. Prema hipotezi o specifičnosti domene, sposobnost nadgledanja direktno ovisi o znanju o domeni. Prema hipotezi o općoj sposobnosti nadgledanja sposobnost je nadgledanja neovisna o specifičnim znanjima i ona je odgovorna za nadgledanje u različitim domenama. Autori su koristili dvije mjere točnosti procjena o izvedbi: diskriminaciju i pristranost. Diskriminacija se odnosi na razliku između prosječnih procjena za netočne odgovore i prosječnih procjena za točne odgovore, transformiranu u vrijednosti između -1 i +1, s time da viši pozitivan rezultat znači pripisivanje viših procjena točnim odgovorima i nižih netočnim odgovorima, a negativan pripisivanje viših procjena netočnim odgovorima i nižih točnim odgovorima. Mjera se pristranosti određuje na temelju razlike između sudionikove prosječne procjene i prosječne izvedbe na svakom testu. Ova se mjera također kreće u rasponu od -1 do +1, s tim da viši rezultat znači pretjeranu samouvjerenost, a niži rezultat znači prenisku samouvjerenost. Pozitivne bi korelacije između mjera diskriminacije kod različitih testova upućivale na opću sposobnost nadgledanja, međutim, dobiveno je da te mjere uglavnom nisu povezane, što je rezultat koji ide u prilog hipotezi o specifičnosti domene. S druge su strane mjere pristranosti za različite testove

međusobno bile povezane, što upozorava na postojanje opće tendencije prema pretjeranoj ili preniskoj samouvjerenosti. Schraw, Potenza i Nebelsick-Gullet (1993) također su uočili tendenciju sudionika da daju konzistentne procjene o izvedbi kroz različite zadatke, bez obzira na stvarnu izvedbu na tim zadacima. Moguće je da takva tendencija proizlazi iz općih očekivanja o vlastitoj izvedbi na testu (npr. iznadprosječnoj, prosječnoj ili ispodprosječnoj).

Schraw (1997) je dobio rezultate koji idu u prilog hipotezi o općoj sposobnosti nadgledanja. Njegovi su sudionici rješavali četiri testa (razumijevanje, leksičke procjene, matematičko zaključivanje i silogizmi) i davali procjene o izvedbi na svakom testu. Ovi su rezultati uspoređeni sa samoprocjenama vlastite općenite sposobnosti nadgledanja prikupljenim pomoću Liste općih strategija nadgledanja (*General Monitoring Strategies Checklist*, GMSC). Dobiveno je da procjene o izvedbi na sva četiri testa koreliraju sa samoprocjenama sposobnosti nadgledanja, što je interpretirano time da se sudionici oslanjaju na vlastito metakognitivno znanje pri procjenjivanju izvedbe u različitim domenama.

Istraživanje koje su proveli Keleman, Frost i Weaver (2000) dovodi u pitanje postojanje opće sposobnosti nadgledanja. Oni su usporedili tri mjere metamemorije (procjene lakoće učenja, procjene u učenju i uvjerenje o znanju) i nadgledanje razumijevanja. Iako su rezultati na zadacima pamćenja i razumijevanja bili korelirani, to nije dobiveno i za točnost metakognitivnih procjena. S druge strane, bile su povezane veličine procjena što opet upućuje na opću tendenciju da se precjenjuje ili podcjenjuje vlastita izvedba. Slične su rezultate dobili i Maki, Shields, Wheeler i Zacchilli (2005), koji su ispitivali povezanost između relativne i apsolutne točnosti procjena o razumijevanju teksta i koji nisu dobili značajnu povezanost između tih mjera. Griffin, Jee i Wiley (2009) također nisu dobili povezanost između relativne točnosti procjena o razumijevanju tekstova iz dviju različitih domena. Na temelju se svih navedenih istraživanja ne može zaključiti da procjene razumijevanja ovise o općoj sposobnosti nadgledanja, odnosno, o općoj sposobnosti da se daju točne procjene o vlastitoj izvedbi, a ne može se čak ni zaključiti da takva opća sposobnost postoji. Međutim, metakognitivne su procjene u različitim domenama, pa tako i kod razumijevanja teksta, pod utjecajem opće tendencije da se podcjenjuje ili precjenjuje vlastita izvedba.

Što se tiče drugoga problema, odnosno povezanosti između točnosti nadgledanja razumijevanja i samoga razumijevanja, u više je istraživanja pokazano da vješti čitatelji posjeduju bolje metakognitivno znanje (npr. Moore, Zabrocky i Commander, 1997) i u skladu je s time pretpostavljeno da će oni biti uspješniji i u kalibraciji razumijevanja, odnosno da će njihove procjene o razumijevanju biti točnije od manje vještih čitatelja.

Maki i Berry (1984, eksperiment 1) podijelili su sudionike u dvije grupe s obzirom na medijan na testu razumijevanja. Sudionici su čitali ekspozitorni tekst i procjenjivali koliko će biti uspješni na testu razumijevanja koji je proveden jedan dan nakon čitanja teksta. Sudionici su općenito pokazali nizak stupanj kalibracije i

pretjeranu samouvjerenost. Sudionici su s rezultatom iznad medijana bili nešto bolji u procjenama od sudionika ispod medijana. Glover (1989) je proveo istraživanje na više od 200 sudionika i usporedio je kalibraciju kod 15 najboljih sudionika na testu razumijevanja i 15 najlošijih. Dobri su čitatelji pokazali znatno višu točnost nadgledanja od loših čitatelja, koji su uz to pokazali i visok stupanj pretjerane uvjerenosti u vlastitu izvedbu na testu razumijevanja.

Iako rezultati ovih istraživanja pokazuju da točnost nadgledanja razumijevanja ovisi o općoj sposobnosti čitanja, nekoliko drugih istraživanja dovodi u pitanje ovu tvrdnju. Tako npr. Pressley i Ghatala (1988) nisu dobili značajnu povezanost između sposobnosti čitanja i točnosti naknadnih procjena. Ovu povezanost nisu dobili ni Commander i Stanwyck (1997). Oni su grupirali sudionike u dobre i loše čitače s obzirom na rezultat na Nelson-Dennyjevu testu čitanja. Sudionici su čitali ekspozitorne odlomke i procjenjivali razumijevanje, odgovarali na pitanja o tekstovima i davali procjene o izvedbi. Ni točnost procjena ni točnost naknadnih procjena nisu bile povezane sa sposobnošću čitanja. Maki, Jonas i Kallod (1994) su mjerili sposobnost razumijevanja Multimedijalnim testom razumijevanja (MMCB; Gernsbacher, Varner i Faust, 1990) i Nelson-Dennyjevim testom. Razumijevanje nije bilo povezano s točnošću procjena nakon čitanja, ali je bilo povezano s točnošću procjena nakon rješavanja zadataka. Maki i sur. (2005) isto tako nisu dobili korelaciju između sposobnosti čitanja i relativne točnosti procjena, ali su dobili negativnu korelaciju između razumijevanja i tendencije prema precjenjivanju vlastite izvedbe.

U skladu se s pregledom istraživanja može zaključiti da je u malo istraživanja do sada ispitan odnos između različitih mjera točnosti nadgledanja razumijevanja, njihova stabilnost u različitim situacijama nadgledanja, kao i njihova povezanost s razumijevanjem. Međutim, ispitivanje takvih problema može imati važne metodološke i teorijske implikacije. Metodološke se implikacije odnose na primjerenost korištenja različitih mjera točnosti nadgledanja, a teorijske na mogućnost izvođenja generalizacija o prirodi procesa nadgledanja na temelju rezultata istraživanja koja koriste određene mjere nadgledanja razumijevanja.

Osnovni je cilj ovoga istraživanja bio ispitati odnose između mjera apsolutne i relativne točnosti nadgledanja u različitim fazama nadgledanja razumijevanja narativnih i ekspozitornih tekstova, kao i povezanosti tih mjera sa samim razumijevanjem. Pri tome je ispitano nekoliko specifičnih problema. Prvo, ispitani su odnosi između relativne točnosti mjerene u različitim uvjetima: prije ispitivanja razumijevanja i nakon njega, za zadatke višestrukoga izbora i za zadatke otvorenoga tipa i s obzirom na vrstu teksta – ekspozitorni i narativni. Prethodna istraživanja potvrđuju da relativna točnost uglavnom pokazuje nisku stabilnost, zbog čega je pretpostavljeno da mjere relativne točnosti u različitim situacijama nadgledanja neće biti međusobno povezane (Griffin, Jee i Wiley, 2009; Kelemen, Frost i Weaver, 2000; Maki i sur., 2005). Drugo, slično je ispitivanje provedeno i za apsolutnu točnost operacionaliziranu kao indeks pristranosti, odnosno, kao

pokazatelj procjenjivanja ili podcjenjivanja vlastite izvedbe. U prethodnim je istraživanjima dobivena veća stabilnost apsolutne točnosti u različitim situacijama nadgledanja i stoga je pretpostavljeno da će različite mjere apsolutne točnosti biti međusobno pozitivno povezane (Kelemen, Frost i Weaver, 2000; Maki i sur., 2005; Schraw i sur., 1995). Treće, ispitane su korelacije između različitih mjera točnosti procjena i razumijevanja. Pretpostavljeno je da relativna točnost neće biti povezana s razumijevanjem te da će apsolutna točnost negativno korelirati s razumijevanjem, odnosno, da će lošija izvedba na testu razumijevanja biti praćena s procjenjivanjem vlastite izvedbe (Maki, Jonas i Kallod, 1994; Maki i sur., 2005).

Metoda

Sudionici

U istraživanju je sudjelovalo 69 studenata (4 studenta i 65 studentica) od prve do treće godine studija psihologije na Filozofskom fakultetu Sveučilišta u Rijeci. Raspon dobi sudionika bio je 18 - 23 godine, prosječna dob sudionika bila je 20.2 godina.

Materijali

Za svakoga su sudionika pripremljene dvije knjižice sa zadacima. Jedna je knjižica sadržavala 6 ekspoziornih, a druga 6 narativnih tekstova, kao i odgovarajuće skale procjena i zadatke za ispitivanje razumijevanja. Svaki je tekst bio otisnut na jednoj stranici, a nakon svakoga teksta bile su još četiri stranice sa skalama procjena i sa zadacima za ispitivanje razumijevanja. Na drugoj su se stranici nalazile skale procjena na kojima je trebalo dati svoje procjene nakon čitanja. Zatim su se na dvije stranice nalazili zadaci za ispitivanje razumijevanja i konačno, na petoj su se stranici nalazile skale procjena koje je trebalo ispuniti nakon rješavanja zadataka za ispitivanje razumijevanja. Unutar svake su knjižice tekstovi bili raspoređeni po slučaju, radi kontrole efekta umora i drugih serijalnih efekata kod rješavanja zadataka. Osim ovih osnovnih knjižica sa zadacima pripremljene su i dvije knjižice sa zadacima za vježbu, koje su sadržavale svaka po jedan tekst za vježbu i odgovarajuće zadatke.

Tekstovi. Svi su sudionici čitali, procjenjivali i odgovarali na pitanja o 6 ekspoziornih i 6 narativnih tekstova. U predispitivanju na malom broju sudionika ($n = 9$) provjereno je da tekstovi ni zadaci nisu prelagani niti preteški. Prosječna je dužina ekspoziornih tekstova bila 371 riječ, narativnih 501 riječ. Narativni su tekstovi bili dulji od ekspoziornih, što je bio slučaj i u istraživanju koje su proveli Weaver i Bryant (1995), zbog toga što se uobičajeno narativni tekstovi smatraju lakšima od ekspoziornih. Tekstovi su bili nešto dulji od tekstova koji su uobičajeno

korišteni u većini istraživanja metarazumijevanja zbog toga da se o svakome tekstu može postaviti više pitanja višestrukoga izbora te pitanja otvorenoga tipa.

Procjene. Za svaki je korišteni tekst od sudionika traženo da naprave dvije procjene nakon čitanja teksta i nakon rješavanja zadataka za ispitivanje razumijevanja. Nakon čitanja su napravili sljedeće procjene:

Procjena rješavanja zadataka višestrukoga izbora. Za svaki su tekst sudionici procjenjivali koliko će zadataka višestrukoga izbora točno riješiti, a raspon je vrijednosti od 0 do 8.

Procjena rješavanja zadatka otvorenoga tipa. Za svaki su tekst sudionici procjenjivali koliko će zadataka otvorenoga tipa točno riješiti, a raspon je vrijednosti od 0 do 4.

Nakon rješavanja zadataka razumijevanja sudionici su napravili te iste dvije procjene, s time što su u ovom uvjetu procjenjivali koliko su zadataka višestrukoga izbora i zadataka otvorenoga tipa točno riješili.

Zadaci za ispitivanje razumijevanja. Za svaki je tekst sastavljeno ukupno 12 pitanja za ispitivanje razumijevanja, od toga 8 pitanja višestrukoga izbora s četiri alternative i 4 pitanja otvorenoga tipa na koje se može odgovoriti jednom do dvije rečenice. Pitanja višestrukoga izbora bodovana su tako da je svakom točno riješenom zadatku dan jedan bod. Pitanja su otvorenoga tipa ocjenjivana na sljedeći način: za svako su pitanje razrađeni kriteriji koji određuju što čini potpuno točan odgovor i što čini djelomično točan odgovor. Potpuno točan odgovor je trebao uključivati sve relevantne informacije koje se u pitanju traže i nije smio uključivati netočne tvrdnje, a potpuno točnom odgovoru dan je jedan bod. Djelomično je točnim odgovorima pripisano pola boda. Odgovor je tretiran kao djelomično točan ako je bio nepotpun ili ako je sadržavao netočne informacije.

Radi provjere objektivnosti kodiranja odgovora na zadacima otvorenoga tipa, drugi je procjenjivač kodirao te odgovore na slučajnom uzorku od 5 sudionika za ekspozitorne i 5 sudionika za narativne tekstove. Za ekspozitorne tekstove dobiveno slaganje iznosi .89, a za narativne .93, što se može smatrati zadovoljavajućim slaganjem.

Mjere točnosti nadgledanja. U analizi su rezultata korištene dvije vrste mjera točnosti nadgledanja. Kao mjera relativne točnosti metakognitivnih procjena korišten je Goodman-Kruskalov gama-koeficijent korelacije. Treba naglasiti da mjere relativne točnosti nisu temeljene na uobičajenim interindividualnim korelacijskim analizama, već je pri izračunu gama-korelacija korištena sljedeća procedura: svaki je sudionik procijenio svaki od šest ekspozitornih tekstova na četiri skale procjena (procjena izvedbe na zadacima višestrukoga izbora i zadacima otvorenoga tipa prije čitanja i nakon njega), a i za svaki je od tih tekstova rješavao dvije vrste zadataka za ispitivanje razumijevanja (zadaci višestrukoga izbora i

zadaci otvorenoga tipa), a isti su podaci dobiveni i za svaki od šest narativnih tekstova. Gama-korelacija se računa za svakoga sudionika, kao korelacija između pojedine procjene (6 podataka, po jedan za svaki tekst) i pojedine mjere razumijevanja (6 podataka, po jedan za svaki tekst), prema sljedećoj formuli:

$$\text{gama} = (C - D) / (C + D),$$

gdje je C broj sukladnih parova tekstova, a D je broj nesukladnih parova tekstova. Tako dobivena gama-korelacija predstavlja individualni indeks relativne točnosti procjene, s obzirom na vrstu procjene i mjeru razumijevanja. S obzirom da je u ovome istraživanju korišteno više skala procjena i više vrsta zadataka za ispitivanje razumijevanja, kao i dvije vrste tekstova, za svakoga je sudionika izračunato više gama-korelacija: procjene rješavanja zadataka višestrukoga izbora i zadataka otvorenoga tipa prije rješavanja i nakon njega uspoređene s odgovarajućim mjerama razumijevanja.

Kao pokazatelj je apsolutne točnosti procjena korišten indeks pristranosti (Schraw i sur., 1995). Indeks pristranosti računat je za procjene rješavanja zadataka višestrukoga izbora za svakog sudionika tako da je izračunata proporcija procjene rješavanja zadataka višestrukoga izbora (npr. ako je sudionik za prvi tekst procijenio da će riješiti 4 zadatka višestrukoga izbora, za drugi 5, i dalje 6, 7, 5 i 6, to daje ukupno 33, odnosno proporciju $33/48 = .69$) i proporcija riješenih zadataka višestrukoga izbora i druga je vrijednost oduzeta od prve (Schraw i sur., 1995). Ovaj je postupak proveden i za procjene nakon čitanja i za procjene nakon rješavanja zadataka. Na sličan je način izračunat indeks pristranosti i za zadatke otvorenoga tipa. Vrijednost indeksa +1.0 značila bi ekstremno precjenjivanje u situaciji kada sudionik procjenjuje da će na svakom tekstu riješiti sve zadatke, a ne riješi ni jedan, dok bi vrijednost od -1.0 značila ekstremno podcjenjivanje. Ako je indeks pristranosti jednak nuli, to bi značilo potpunu apsolutnu točnost procjena, jer za svaki tekst procijenjen broj točnih odgovora odgovara stvarnom broju točnih odgovora. Međutim, ova interpretacija nije potpuno točna jer je moguće da se sudionik na nekim tekstovima precjenjuje, a na nekim podcjenjuje, a te se vrijednosti mogu poništiti jer se u izračunu indeksa pristranosti vodi računa o predznaku. Ipak, ovako je izračunat indeks pristranosti dobar pokazatelj sustavne tendencije sudionika prema precjenjivanju ili podcjenjivanju.

Postupak

Istraživanje se sastojalo od dva dijela: ispitivanja metarazumijevanja ekspozitornih tekstova i ispitivanja metarazumijevanja narativnih tekstova. Polovica je sudionika određenih po slučaju prvo pristupila ispitivanju metarazumijevanja ekspozitornih tekstova, a polovica sudionika narativnih tekstova. Ispitivanje je metarazumijevanja druge vrste tekstova provedeno unutar dva tjedna (minimalno 10 dana) od ispitivanja metarazumijevanja prve vrste tekstova. Ispitivanje je bilo grupno, a veličina je grupe varirala od 5 do 14 sudionika.

Na početku je sudionicima rečeno koji je cilj ispitivanja i kratko je opisan način ispitivanja. Naglašeno je da tekstove trebaju čitati s razumijevanjem, svojim tempom. Nije bilo dopušteno vraćati se na tekst kada se započelo s procjenjivanjem, ni vraćati se na riješene zadatke kada bi se započelo s procjenjivanjem rješavanja zadataka, ni vraćati se na prethodne tekstove. Nakon rješavanja zadatka za vježbu pristupilo se glavnom dijelu ispitivanja. Sudionici su pročitali tekst, napravili tražene procjene nakon čitanja, rješavali zadatke razumijevanja i napravili procjene nakon rješavanja zadatka. Kada bi završili sa svim zadacima koji se odnose na tekst, prelazili bi na sljedeći tekst. Ispitivanje nije bilo vremenski ograničeno, a kako su sudionici radili svojim tempom, bilo im je potrebno različito vrijeme za rješavanje zadataka. Maksimalno je vrijeme potrebno za rješavanje bilo oko 90 minuta, ali većina je glavni dio eksperimentalne procedure (čitanje 6 tekstova i rješavanje zadataka) u oba dijela istraživanja (ekspozitorni i narativni tekstovi) obavila za 60 – 70 minuta.

Rezultati

U Tablici 1. su prikazane prosječne procjene nakon čitanja i nakon rješavanja zadataka, kao i rezultati na zadacima za ispitivanje razumijevanja. S obzirom da ovdje nije bio cilj raditi specifične usporedbe s obzirom na pojedine tekstove, tekstovi nisu međusobno uspoređivani, ali se s obzirom na prosječne rezultate može zaključiti da je težina tekstova i zadataka bila zadovoljavajuća (Weaver, 1990). Posebno treba istaknuti da zadaci nisu bili preteški, jer u toj situaciji može doći do niže točnosti procjenjivanja (Weaver i Bryant, 1995).

Tablica 1. Prosječne procjene i izvedba na zadacima za ispitivanje razumijevanja

	Tip zadatka	Vrsta teksta			
		Ekspozitorni		Narativni	
		M	SD	M	SD
Procjene nakon čitanja	Višestruki izbor	5.05	0.96	5.24	0.91
	Otvoreni tip	1.97	0.50	2.07	0.43
Procjene nakon rješavanja	Višestruki izbor	4.87	1.05	5.56	0.85
	Otvoreni tip	1.90	0.62	2.40	0.58
Razumijevanje	Višestruki izbor	34.13	4.35	35.45	4.44
	Otvoreni tip	12.20	3.38	16.63	3.04

U analizi su rezultata korištene tri grupe varijabli: mjere relativne točnosti, mjere apsolutne točnosti i mjere razumijevanja. Kao mjera je relativne točnosti korištena gama-korelacija između pojedinih procjena i izvedbe na zadacima razumijevanja, a kao mjera apsolutne točnosti razlika između procjena i izvedbe na zadacima razumijevanja. U Tablici 2. prikazane su prosječne vrijednosti mjera relativne i apsolutne točnosti po specifičnim uvjetima.

Što se tiče relativne točnosti procjena nakon čitanja, može se reći da je ta točnost slična onoj dobivenoj u većini istraživanja u kojima nisu korišteni posebni postupci s ciljem dobivanja više točnosti procjena, gdje se prosječna gama-korelacija između procjena i izvedbe obično kreće u rasponu od 0 do .30. Što se tiče kalibracije izvedbe, odnosno, točnosti procjena nakon rješavanja zadataka razumijevanja, također je dobivena relativna točnost procjena (raspon gama od .35 do .72) slična onoj u drugim istraživanjima (Pierce i Smith, 2001). Što se tiče apsolutne točnosti procjena, u prosjeku, na svim su mjerama sudionici pokazali blagu tendenciju prema podcjenjivanju, iako je znatan dio sudionika pokazao i precjenjivanje.

Tablica 2. Prosječne gama-korelacije (relativna točnosti) i indeksi pristranosti (apsolutna točnost) u različitim situacijama nadgledanja

Vrsta teksta	Vrijeme procjena	Tip zadatka	Relativna točnost ^a		Apsolutna točnost ^b	
			M	SD	M	SD
Ekspozitorni	Nakon čitanja	Višestruki izbora	0.31	0.50	-0.08	0.11
		Otvoreni tip	0.45	0.49	-0.02	0.15
	Nakon rješavanja	Višestruki izbora	0.50	0.43	-0.10	0.11
		Otvoreni tip	0.72	0.35	-0.04	0.14
Narativni	Nakon čitanja	Višestruki izbora	0.32	0.50	-0.08	0.13
		Otvoreni tip	0.29	0.59	-0.18	0.13
	Nakon rješavanja	Višestruki izbora	0.52	0.51	-0.04	0.11
		Otvoreni tip	0.56	0.39	-0.09	0.11

^aGama-korelacije; ^bIndeksi pristranosti

Ispitane su povezanosti između mjera relativne točnosti procjena u različitim uvjetima, a dobivene su korelacije prikazane u Tablici 3.

Tablica 3. Korelacije između mjera relativne točnosti u različitim situacijama nadgledanja

Vrsta teksta	Vrijeme procjena	Tip zadatka							
			1	2	3	4	5	6	7
Ekspozitorni	Nakon čitanja	1. VI							
		2. OT	.16						
	Nakon rješavanja	3. VI	.53**	-.09					
		4. OT	-.29*	.04	-.20				
Narativni	Nakon čitanja	5. VI	.07	.01	-.04	-.10			
		6. OT	.11	-.22	.20	.01	.04		
	Nakon rješavanja	7. VI	.01	.01	.00	.03	-.14	.35**	
		8. OT	.02	-.09	.08	-.17	-.04	.47**	.07

* $p < .05$; ** $p < .01$; Tip zadatka: VI - zadaci višestrukoga izbora, OT - zadaci otvorenoga tipa

S obzirom na svrhu ove analize valja istaknuti dva ključna nalaza. Prvo, ni jedna mjera relativne točnosti za ekspozitorne tekstove nije značajno povezana s mjerama relativne točnosti za narativne tekstove. Drugo, većina korelacija između relativne točnosti procjena unutar istih tekstova, ali za različite vrste zadataka i faze procjenjivanja također nije značajna. Iznimka je od toga pozitivna korelacija između točnosti procjena za zadatke višestrukoga izbora kod ekspozitornih tekstova prije i nakon čitanja ($r = .53$), a za narativne tekstove korelacija između točnosti procjena za zadatke otvorenoga tipa prije i nakon čitanja ($r = .47$) te korelacija između točnosti procjena za zadatke otvorenoga tipa prije čitanja i zadatke višestrukoga izbora nakon rješavanja ($r = .35$). Zaključno, ovi rezultati upozoravaju na nizak stupanj stabilnosti relativne točnosti procjenjivanja, osobito između procjena na različitim vrstama tekstova. Međutim, relativna točnost ipak pokazuje određen stupanj stabilnosti između različitih faza procjenjivanja na istim tekstovima i istom tipu zadataka za ispitivanje razumijevanja.

Dalje su analizirane korelacije između mjera apsolutne točnosti u različitim uvjetima, a dobiveni su rezultati prikazani u Tablici 4.

Tablica 4. Korelacije između mjera apsolutne točnosti u različitim situacijama nadgledanja

Vrsta teksta	Vrijeme procjena	Tip zadatka	1	2	3	4	5	6	7
Ekspozitorni	Nakon čitanja	1. VI							
		2. OT	.46**						
	Nakon rješavanja	3. VI	.81**	.36**					
		4. OT	.36**	.67**	.54**				
Narativni	Nakon čitanja	5. VI	.53**	.34**	.28*	.18			
		6. OT	.21	.47**	.20	.36**	.50**		
	Nakon rješavanja	7. VI	.44**	.23	.32**	.22	.76**	.31*	
		8. OT	.19	.28*	.16	.29*	.45**	.64**	.48**

* $p < .05$; ** $p < .01$; Tip zadatka: VI – zadaci višestrukoga izbora, OT – zadaci otvorenoga tipa

Od 28 koeficijenata korelacije, 21 je značajan i pozitivan, a preostalih 7 veći su od nule, ali nisu statistički značajni. Ovakav obrazac korelacija upućuje na stabilnost indeksa pristranosti kroz različite zadatke razumijevanja, vrste tekstova i faze procjenjivanja, iako su korelacije više za indekse pristranosti kod istog tipa zadatka i istog tipa teksta, nego kod različitih tipova teksta i zadataka. Prema tome, postoji općenita tendencija prema podcjenjivanju ili precjenjivanju, koja se može pokazati u različitim situacijama procjenjivanja. Slične su rezultate dobili Kelemen, Frost i Weaver (2000), Maki i sur. (2005) te Schraw i sur. (1995). Apsolutna je točnost procjena, u smislu tendencije prema podcjenjivanju i precjenjivanju, stabilniji aspekt nadgledanja razumijevanja od relativne točnosti procjena, koja u znatno većoj mjeri varira s obzirom na vrstu zadatka.

Nadalje, veličina se korelacija između indeksa pristranosti razlikuje s obzirom na fazu procjenjivanja, vrstu teksta i vrstu zadatka za ispitivanje razumijevanja. Najviše su korelacije između indeksa pristranosti za procjene na istim vrstama teksta i istom tipu zadatka, ali u različitim fazama procjenjivanja. Niže su korelacije između indeksa pristranosti kod procjena koje se odnose na iste tekstove i istu fazu procjenjivanja, ali na različite zadatke razumijevanja. Konačno, najniže su korelacije između indeksa pristranosti za različite vrste tekstova. Ovi rezultati pokazuju da je i apsolutna točnost procjena u određenoj mjeri pod utjecajem specifičnosti situacije procjenjivanja, ali u znatno manjoj mjeri nego relativna točnost.

Ni jedan koeficijent korelacije između mjera relativne točnosti i mjera apsolutne točnosti nije statistički značajan. Takvi su nalazi dobiveni i u drugim istraživanjima (Maki i sur., 2005; Schraw i sur., 1995). Može se pretpostaviti da relativna i apsolutna točnost metakognitivnih procjena uključuje različite procese (Maki i sur., 2005). Apsolutna je točnost vjerojatno pod utjecajem generalizirane percepcije o vlastitim sposobnostima (Moore, Lin-Agler i Zabrocky, 2005), dok relativna točnost ovisi o specifičnim procesima nadgledanja uključenim kod izvođenja specifičnih zadataka.

Konačno, analizirane su i korelacije između mjera apsolutne i relativne točnosti i razumijevanja, a dobiveni su rezultati prikazani u Tablicama 5. i 6.

Tablica 5. Korelacije između relativne točnosti procjena i razumijevanja

		Ekspozitorni tekstovi				Narativni tekstovi			
		Nakon čitanja		Nakon rješavanja		Nakon čitanja		Nakon rješavanja	
Vrsta teksta	Tip zadatka	VI	OT	VI	OT	VI	OT	VI	OT
Ekspozitorni	VI	.04	-.02	.08	-.04	.09	-.23	-.10	-.17
	OT	.02	.07	.02	.05	.02	-.16	-.08	-.09
Narativni	VI	.11	-.03	.12	-.15	-.02	-.18	-.08	-.10
	OT	.05	.11	.08	-.03	-.19	-.18	-.14	.04

Tip zadatka: VI – zadaci višestrukoga izbora, OT – zadaci otvorenoga tipa

Tablica 6. Korelacije između apsolutne točnosti procjena i razumijevanja

Vrsta teksta	Tip zadatka	Ekspozitorni tekstovi				Narativni tekstovi			
		Nakon čitanja		Nakon rješavanja		Nakon čitanja		Nakon rješavanja	
		VI	OT	VI	OT	VI	OT	VI	OT
Ekspozitorni	VI	-.30*	-.33**	-.14	-.08	-.15	-.30*	-.12	.10
	OT	-.02	-.62**	.10	-.37**	-.16	-.42**	-.10	-.02
Narativni	VI	-.15	-.16	-.00	-.03	-.51**	-.42**	-.46**	-.18
	OT	.02	-.19	.01	-.15	-.20	-.67**	-.08	-.28*

* $p < .05$; ** $p < .01$; Tip zadatka: VI – zadaci višestrukoga izbora, OT – zadaci otvorenoga tipa

Ni jedna mjera relativne točnosti nije značajno povezana s mjerama razumijevanja. Prema tome, bolja izvedba na zadacima razumijevanja nije povezana i s višom relativnom točnošću procjenjivanja. Maki, Jonas i Kallod (1994) također nisu dobili značajnu povezanost između razumijevanja i relativne točnosti nakon čitanja, ali su dobili značajne pozitivne korelacije između mjera razumijevanja i točnosti procjena nakon rješavanja zadataka. Pressley i Ghatala (1990) također navode više istraživanja u kojima nije dobivena takva povezanost, a povezanost između verbalnih sposobnosti i relativne točnosti procjena nisu dobili ni Maki i sur. (2005). Svi ovi nalazi pokazuju da, suprotno očekivanjima, točnost nadgledanja određena kao relativna točnost procjena, nije povezana s razumijevanjem, odnosno, da bolji čitatelji nisu ujedno i nužno bolji u nadgledanju razumijevanja.

S druge strane, iako sve mjere razumijevanja nisu značajno povezane sa svim mjerama apsolutne točnosti, može se uočiti jedna pravilnost u korelacijama: značajna je negativna korelacija između pojedinih mjera razumijevanja i odgovarajućih pokazatelja apsolutne točnosti. Npr. koeficijenti korelacije između izvedbe na zadacima višestrukoga izbora kod narativnoga teksta i dva indeksa pristranosti za zadatke višestrukoga izbora kod narativnih tekstova iznose $-.51$ i $-.46$. Slično je dobiveno i za preostale četiri mjere razumijevanja, osim za korelaciju između zadataka višestrukoga izbora kod ekspozitornih tekstova i indeksa pristranosti za procjene nakon rješavanja zadataka, koja je negativna, ali nije značajna. Većina je preostalih koeficijenta korelacija negativna, i kada nisu značajni. Najveće negativne korelacije između razumijevanja i indeksa pristranosti dobivene su za zadatke otvorenoga tipa nakon čitanja.

Slične su rezultate dobili Schraw i sur. (1995, eksperiment 2) i Maki i sur. (2005). Prema tome, sudionici koji pokazuju lošiju izvedbu na određenom zadatku pokazuju tendenciju prema precjenjivanju. Međutim, na temelju se postojećih rezultata ne može generalizirati da će niža sposobnost čitanja općenito biti povezana s precjenjivanjem, iako takva mogućnost svakako postoji.

Rasprava

Najvažniji zaključak koji slijedi iz ovoga istraživanja, a i nekoliko drugih istraživanja (Kelemen, Frost i Weaver, 2000; Maki i sur., 2005), je da se relativna i apsolutna točnost sustavno razlikuju u nekoliko važnih obilježja.

Prvo, mjere relativne i apsolutne točnosti međusobno nisu povezane, a takve su rezultate dobili Kelemen, Frost i Weaver (2000) i Maki i sur. (2005). Takav izostanak povezanosti upućuje na to da su relativna i apsolutna točnost različiti i prilično neovisni aspekti nadgledanja razumijevanja. U skladu s interpretacijom koju daju Maki i sur. (2005) možemo pretpostaviti da apsolutna točnost ovisi u prvom redu o tendenciji da se vlastita izvedba precjenjuje ili podcjenjuje, što odražava općenitu percepciju vlastitih sposobnosti. S druge strane, relativna točnost ovisi o stupnju u kojem varijabilitet u procjenama između pojedinih tekstova odgovara varijabilitetu u izvedbi na testu razumijevanja. Takva povezanost podrazumijeva procjenjivanje na temelju specifičnih iskustava s pojedinim tekstovima.

Drugo, ako pod stabilnošću podrazumijevamo pozitivnu povezanost između točnosti metakognitivnoga nadgledanja u različitim situacijama nadgledanja, relativna i apsolutna točnost pokazuju različiti stupanj stabilnosti kroz različite situacije nadgledanja, ovisno o tipu teksta, mjeri razumijevanja i fazi procjenjivanja. Općenito, može se reći da apsolutna točnost pokazuje veći stupanj stabilnosti. Npr. ni jedna od četiri mjere relativne točnosti za ekspozitorne tekstove ne korelira pozitivno s mjerama relativne točnosti za narativne tekstove. S druge strane, mjere su apsolutne točnosti unutar istoga tipa teksta međusobno umjereno do visoko pozitivno povezane, dok mjere apsolutne točnosti između različitih vrsta tekstova pokazuju blagu do umjerenu pozitivnu povezanost. Ipak, može se reći da što su sličnije situacije nadgledanja, veća je vjerojatnost i da će se točnost nadgledanja podudarati, što vrijedi i za apsolutnu i relativnu točnost.

Što je izvor takvih razlika u stabilnosti različitih mjera točnosti nadgledanja? Stabilnost apsolutne točnosti proizlazi iz procjenjivanja na temelju percepcije vlastitih sposobnosti, koja ne varira previše između različitih zadataka (Maki i sur., 2005). Moore, Lin-Agler i Zabrocky (2005) smatraju da takva općenita percepcija proizlazi iz prethodnih iskustava sa sličnim problemskim situacijama i da ne varira previše pod utjecajem specifičnih situacijskih čimbenika. Što se tiče niske stabilnosti relativne točnosti, ona proizlazi iz drugih čimbenika. Jedna je mogućnost da relativna točnost određena kao gama-korelacija između procjene i izvedbe ima nisku pouzdanost. Na primjer, Kelemen, Frost i Weaver (2000) dobili su da je test-retest pouzdanost takvih mjera vrlo niska, zbog čega može proizlaziti i niska stabilnost takvih mjera. I drugi bi metodološki čimbenici mogli pridonositi niskoj stabilnosti, npr. mali broj zadataka za ispitivanje razumijevanja (Lin i Zabrocky, 1998). Međutim, takve interpretacije ne mogu objasniti stabilne razlike u relativnoj točnosti koje se sustavno dobivaju između različitih eksperimentalnih situacija, npr.

između procjenjivanja nakon čitanja i nakon rješavanja zadataka (Pierce i Smith, 2001). Štoviše, otkriven je niz čimbenika koji utječu na relativnu točnost procjenjivanja (za pregled vidjeti Lin i Zabrocky, 1998; Bajšanski, 2008). Osim ovakvih metodoloških interpretacija niske stabilnosti relativne točnosti nadgledanja može se pretpostaviti da ne postoji opća sposobnost nadgledanja, koja će se manifestirati u različitim situacijama nadgledanja, ili da takva sposobnost uključuje različite komponente, zbog čega između ostaloga visoka točnost nadgledanja razumijevanja narativnih tekstova neće biti povezana s visokom točnošću kod ekspozitornih tekstova.

Treće, relativna i apsolutna točnost pokazuju različite obrasce povezanosti s razumijevanjem. Dok ni jedna mjera relativne točnosti nije pozitivno povezana s razumijevanjem, mjere apsolutne točnosti negativno koreliraju s razumijevanjem, a takvi su obrasci povezanosti dobiveni i u drugim istraživanjima (Maki, Jonas i Kallod, 1994; Maki i sur., 2005, Lin, Moore i Zabrocky 2001; Pressley i Ghatala 1988). Ovdje to znači da će lošija izvedba na zadacima za ispitivanje razumijevanja biti povezana s precjenjivanjem te izvedbe i pretjeranom samouvjerenošću u točnost odgovora. S druge strane, za višu relativnu točnost možemo zaključiti da nije manifestacija bolje sposobnosti razumijevanja, već ovisi o drugim čimbenicima.

Sve ovo pokazuje da su relativna i apsolutna točnost različiti i prilično neovisni aspekti nadgledanja razumijevanja. Možda je najvažnija implikacija toga nalaza metodološka. U različitim su istraživanjima korištene različite mjere relativne i apsolutne točnosti, u većini istraživanja jedna mjera, a na temelju se rezultata tih istraživanja ponekad izvode općenite tvrdnje o procesima metakognitivnoga nadgledanja razumijevanja. Međutim, ako se te različite mjere razlikuju na ovakav sustavan način, mogućnost je izvođenja generalizacija vrlo ograničena. Najbolja će strategija zbog toga biti korištenje različitih mjera točnosti procjena i ispitivanje njihovih odnosa, a ne generaliziranje na temelju jedne mjere.

Teorijska je implikacija dobivenih rezultata, kao i rezultata drugih spomenutih istraživanja, potrebna za preciziranjem odnosa između apsolutne i relativne točnosti procjena. Pitanje je, kada sudionici istraživanja predviđaju ili procjenjuju vlastitu izvedbu i kada te podatke uspoređujemo sa stvarnom izvedbom, što zapravo mjerimo. Ovisno o odabranome postupku usporedbe rezultati mogu biti različiti. Prema tome, potrebno je objasniti o čemu ovisi procjena koju daju sudionici kada ih se pita koliko će zadataka točno riješiti ili u kojoj su mjeri razumjeli tekst. Veličina će procjena ovisiti o općoj percepciji vlastitih sposobnosti i ona će određivati apsolutnu točnost procjena, odnosno sustavno precjenjivanje ili podcjenjivanje vlastite izvedbe. Međutim, sama veličina nije važna za relativnu točnost. Za viši stupanj relativne točnosti potrebno je da razlike u procjenama odgovaraju razlikama u razumijevanju, odnosno da postoji korelacija između tih dviju varijabli. Da bi postojala korelacija između tih varijabli, potrebno je općenitu procjenu sposobnosti modificirati u skladu sa specifičnim iskustvima sa situacijom

čitanja i razumijevanja ili rješavanja zadataka. Ipak, ovo je površna analiza koja samo upozorava da su potrebna daljnja istraživanja odnosa između relativne i apsolutne točnosti nadgledanja.

Literatura

- Bajšanski, I. (2008). *Metakognitivne procjene i razumijevanje teksta*. Doktorska disertacija. Zagreb: Filozofski fakultet Sveučilišta u Zagrebu.
- Commander, N.E. i Stanwyck, D.J. (1997). Illusion of knowing in adult readers: Effects of reading skill and passage length. *Contemporary Educational Psychology*, 22, 39-52.
- Dunlosky, J., Rawson, K.A. i Middleton, E.L. (2005). What constrains the accuracy of metacomprehension judgments? Testing the transfer-appropriate-monitoring and accessibility hypotheses. *Journal of Memory and Language*, 52, 551-565.
- Epstein, W., Glenberg, A.M. i Bradley, M.M. (1984). Coactivation and comprehension: Contribution of text variables to the illusion of knowing. *Memory and Cognition*, 12, 355-360.
- Fischhoff, B., Slovic, P. i Lichtenstein, S. (1977). Knowing with certainty: The appropriateness of extreme confidence. *Journal of Experimental Psychology: Human Perception and Performance*, 3, 552-564.
- Gernsbacher, M.A., Varner, K.R. i Faust, M.E. (1990). Investigating differences in general comprehension skill. *Journal of Experimental Psychology: Language, Memory, and Cognition*, 16, 430-445.
- Glenberg, A.M. i Epstein, W. (1985). Calibration of comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 11(4), 702-718.
- Glenberg, A.M., Sanocki, T., Epstein, W. i Morris, C. (1987). Enhancing calibration of comprehension. *Journal of Experimental Psychology: General*, 116(2), 119-136.
- Glenberg, A.M., Wilkinson, A. i Epstein, W. (1982). The illusion of knowing: Failure in the self-assessment of comprehension. *Memory and Cognition*, 10, 597-602.
- Glover, J.A. (1989). Improving readers' estimates of learning from text: The role of inserted questions. *Reading Research Quarterly*, 28, 68-75.
- Griffin, T.D., Jee, B.D. i Wiley, J. (2009). The effects of domain knowledge on metacomprehension accuracy. *Memory and Cognition*, 37, 1001-1013.
- Keleman, W.L., Frost, P.J. i Weaver, C.A. III. (2000). Individual differences in metacognition: Evidence against a general metacognitive ability. *Memory and Cognition*, 28, 92-107.
- Lin, L., Moore, D. i Zabucky, K.M. (2001). An assessment of students' calibration of comprehension and calibration of performance using multiple measures. *Reading Psychology*, 22, 111-128.
- Lin, L. i Zabucky, K.M. (1998). Calibration of comprehension: Research and implications for education and instruction. *Contemporary Educational Psychology*, 23, 345-391.

- Maki, R.H. i Berry, S.L. (1984). Metacomprehension of text material. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 10, 663-679.
- Maki, R.H., Jonas, D. i Kallod, M. (1994). The relationship between comprehension and metacomprehension ability. *Psychonomic Bulletin and Review*, 1, 126-138.
- Maki, R.H. i Serra, M. (1992a). The basis of test predictions for text material. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 18(1), 116-126.
- Maki, R.H. i Serra, M. (1992b). Role of practice tests in the accuracy of test predictions on text material. *Journal of Educational Psychology*, 84(2), 200-210.
- Maki, R.H., Shields, M., Wheeler, A.E. i Zacchilli, T.L. (2005). Individual differences in absolute and relative metacomprehension accuracy. *Journal of Educational Psychology*, 97, 723-731.
- Moore, D., Lin-Agler, L.M. i Zabucky, K. (2005). A source of metacomprehension inaccuracy. *Reading Psychology*, 26, 251-265.
- Moore, D., Zabucky, K. i Commander, N.E. (1997). Validation of the metacomprehension scale. *Contemporary Educational Psychology*, 22, 457-471.
- Nelson, T.O. (1984). A comparison of current measures of the accuracy of feeling-of-knowing predictions. *Psychological Bulletin*, 95, 109-133.
- Nelson, T.O. i Narens, L. (1990). Metamemory: A theoretical framework and some new findings. U: G.H. Bower (Ur.), *The psychology of learning and motivation: Advances in research and theory* (str. 125-173). New York: Academic Press.
- Pierce, B.H. i Smith, S.M. (2001). The postdiction superiority effect in metacomprehension of text. *Memory and Cognition*, 29, 62-67.
- Pressley, M. i Ghatala, E.S. (1988). Delusions about performance on multiple-choice comprehension tests. *Reading Research Quarterly*, 23(4), 454-464.
- Pressley, M. i Ghatala, E.S. (1990). Self-regulated learning: Monitoring learning from text. *Educational Psychologist*, 25(1), 19-33.
- Schraw, G. (1997). The effect of generalized metacognitive knowledge on test performance and confidence judgments. *The Journal of Experimental Education*, 65(2), 135-146.
- Schraw, G., Dunkle, M.E., Bendixen, L.D. i Roedel, T.D. (1995). Does a general monitoring skill exist? *Journal of Educational Psychology*, 87, 433-444.
- Schraw, G., Potenza, M.T. i Nebelsick-Gullet, L. (1993). Constraints on the calibration of performance. *Contemporary Educational Psychology*, 18, 455-463.
- Weaver, C.A. III. (1990). Constraining factors in calibration of comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 16(2), 214-222.
- Weaver, C.A. III. i Bryant, D.S. (1995). Monitoring of comprehension: The role of text difficulty in metamemory for narrative and expository text. *Memory and Cognition*, 23(1), 12-22.

Monitoring Narrative and Expository Text Comprehension: The Relationship between Absolute and Relative Metacomprehension Accuracy

Abstract

Metacomprehension accuracy can be conceptualized in two different ways: relative accuracy refers to the correlation between judgments of comprehension and comprehension performance, and absolute accuracy refers to the difference between judgments and performance. The aim of this study was to explore the relations between measures of relative and absolute metacomprehension accuracy across different monitoring tasks and measures of text comprehension. Psychology students ($N = 69$) participated in the study. They read 6 expository and 6 narrative texts. After reading a text they predicted performance on specific types of comprehension questions. Comprehension of each text was assessed by 8 multiple choice questions and 4 open-ended questions. After taking a test participants rated their performance. The gamma correlation coefficient between judgments and comprehension performance was used as an index of relative judgment accuracy. Judgment bias (a difference between performance judgment and text comprehension test result) was used as an index of absolute judgment accuracy. Correlation analysis revealed different patterns of correlations for measures of relative and absolute accuracy. Correlations between measures of relative accuracy of judgments in different conditions (narrative vs. expository text, multiple choice vs. open-ended questions and after reading vs. after taking a test) were much lower than correlations between measures of absolute accuracy. Relative accuracy was not related to text comprehension. Absolute accuracy was related to comprehension: poor performance on the comprehension test was related to overconfidence. Overall, the results indicate that relative and absolute accuracy differ in several aspects and that they involve different metacognitive processes.

Keywords: metacognition, metacomprehension, comprehension monitoring, judgment accuracy

Primljeno: 26.11.2010.

