

FOR PUBLICATION
UNITED STATES COURT OF APPEALS
FOR THE NINTH CIRCUIT

ALONZO DEON JOHNSON; DARRYL THOMPSON, <i>Petitioners-Appellants,</i> v. CLAUDE E. FINN, WARDEN; ATTORNEY GENERAL FOR THE STATE OF CALIFORNIA; TOM L. CAREY, <i>Respondents-Appellees.</i>
--

No. 10-15641
D.C. Nos.
2:03-cv-02063-
JAM-JFM
2:04-cv-02208-
JAM-JFM
OPINION

Appeal from the United States District Court
for the Eastern District of California
John A. Mendez, District Judge, Presiding

Argued and Submitted
October 14, 2011—San Francisco, California

Filed December 8, 2011

Before: Betty B. Fletcher, Stephen Reinhardt, and
A. Wallace Tashima, Circuit Judges.

Opinion by Judge Reinhardt

COUNSEL

Daniel J. Broderick, Federal Defender, Sacramento, California; David M. Porter (argued), Assistant Federal Defender, Sacramento, California; for the petitioners-appellants.

Kamala D. Harris, Attorney General of California, Sacramento, California; Michael P. Farrell, Senior Assistant Attorney General, Sacramento, California; Catherine Chatman, Supervising Deputy Attorney General, Sacramento, California; R. Todd Marshall (argued), Deputy Attorney General, Sacramento, California; for the respondents-appellees.

OPINION

REINHARDT, Circuit Judge:

Alonzo Deon Johnson and Darrell Thompson, California state prisoners, challenge the prosecution's use of peremptory strikes to exclude black jurors in their trial. A magistrate judge, after holding an evidentiary hearing at which the prosecutor testified, found that he had purposefully discriminated on the basis of race in exercising a peremptory strike against one of the black jurors. The district judge, without holding a new evidentiary hearing, rejected the magistrate judge's finding as to the prosecutor's lack of credibility in asserting race-neutral reasons for having stricken the juror. In doing so, the district judge denied Johnson and Thompson the process that they were constitutionally due.

We hold that the rule of *United States v. Ridgway*, 300 F.3d 1153 (9th Cir. 2002), extends to determinations by a magistrate judge as to the credibility of a prosecutor's testimony at the second and third steps of the inquiry required by *Batson v. Kentucky*, 476 U.S. 79 (1986). In *Ridgway*, we held that the Due Process Clause required "that a district court . . . conduct its own evidentiary hearing before rejecting a magistrate judge's credibility findings made after a hearing on a motion to suppress." 300 F.3d at 1154. As in *Ridgway*, an in-person evaluation of a witness's demeanor—here, that of the prosecutor—is essential to the kind of determination that the district judge was required to make: "In the typical peremptory challenge inquiry, the decisive question will be whether counsel's race-neutral explanation for a peremptory challenge should be believed. There will seldom be much evidence bearing on that issue, and the best evidence often will be the demeanor of the attorney who exercises the challenge." *Hernandez v. New York*, 500 U.S. 352, 365 (1991). The district judge erred by declining the opportunity to observe the trial prosecutor's demeanor before rejecting the magistrate judge's adverse credibility finding.

We therefore vacate the district court's denial of the writ of habeas corpus and remand for the district judge either to accept the magistrate judge's credibility finding or to conduct a new evidentiary hearing. We retain jurisdiction over any appeal from the district court's judgment.

I

In 2000, Johnson and Thompson were tried together for murder and other charges in the death of Rafael Palacios. They were acquitted of murder but convicted of shooting at an occupied motor vehicle and, in Thompson's case, of willfully participating in a street gang and being a felon in possession of a firearm. Several sentence enhancements were found to apply in each case.

During the jury selection phase of their trial, Johnson and Thompson raised objections under *Batson* and its state-law cognate, *People v. Wheeler*, 22 Cal. 3d 258 (1978), to the prosecution's use of peremptory challenges against three black jurors: W.J., E.G., and W.T. The trial court found in each case that Johnson and Thompson "had failed to make a prima facie showing that the prosecutor had an invidious basis for the peremptory challenge."

After exhausting his remedies in state court, including an appeal before the intermediate state appellate court and a petition for review that the state supreme court declined to hear, Johnson filed a timely petition for a writ of habeas corpus in the U.S. District Court for the Eastern District of California. Thompson did the same in the Northern District of California. Thompson's case was transferred to the Eastern District, the state filed answers to both petitions, and the district court deemed the cases related.

Magistrate Judge John F. Moulds issued an order concluding that the California Court of Appeal had applied an incorrect legal standard in determining whether Johnson and

Thompson had established a prima facie case of racial discrimination. The magistrate judge therefore determined that he would evaluate Johnson and Thompson's *Batson* claim de novo, without affording deference under the Anti-Terrorism and Effective Death Penalty Act (AEDPA). The magistrate judge found that Johnson and Thompson had made a prima facie showing of racial discrimination as to each of the three black jurors whose strikes were at issue. Recognizing that under *Batson*, "the burden shifts to the state to explain the racial exclusion by offering permissible race-neutral justifications for his strikes," the magistrate judge ordered an evidentiary hearing, as the state had "never been required to present evidence of the prosecutor's actual, non-discriminatory reasons for striking the three black jurors."

After hearing testimony from the trial prosecutor, the magistrate judge issued a forty-three-page report of findings and recommendations. The finding that concerns us here is the magistrate judge's determination that the prosecutor's asserted race-neutral reasons for striking W.J. were not his genuine reasons for doing so. Upon conducting a thorough comparative juror analysis, the magistrate judge concluded that "[a] comparison between [W.J.] and . . . other jurors fatally undermines the credibility of the prosecutor's stated justification for excusing [W.J.] and demonstrates that [W.J.'s] youth, marital status, residence and poor spelling"—all reasons that the prosecutor had given—"could not have genuinely motivated the prosecutor to strike him." The magistrate judge also found that "the prosecutor's failure to ask follow-up voir dire in an effort to clear up his alleged concerns[] suggests he made up nonracial reasons to strike [W.J.]" The magistrate judge therefore found that the prosecutor's "stated reasons for excluding [W.J.] were a pretext for eliminating him from the jury on account of his race"—in other words, that the prosecutor's testimony as to the strike of W.J. was not credible. The magistrate judge found that the prosecutor had *not* discriminated in striking the other two black jurors, E.G. and W.T.

The district judge, in a four-page order, upheld the magistrate judge's findings and recommendations—including those concerning the inapplicability of AEDPA deference—except for the determination that the prosecutor's asserted reasons for striking W.J. were pretextual. The district judge found that Johnson and Thompson did not show “that the totality of circumstances raises an inference that the strike was motivated by race.” He found that the prosecutor “put forward evidence of legitimate, race-neutral reasons for exercising a peremptory challenge against” W.J. and that Johnson and Thompson failed to “prove purposeful racial discrimination by the prosecutor.” In short, the district judge rejected the magistrate judge's finding as to the prosecutor's lack of credibility. Whereas the magistrate judge found that the prosecutor's asserted reasons were not his actual reasons for striking W.J., the district judge found that the prosecutor struck W.J. for “legitimate, race-neutral reasons.” This appeal followed.

II

Before considering whether the district judge was required to hold a new evidentiary hearing in order to reject the credibility determination of the magistrate judge, we must address two threshold questions as to whether it was necessary to hold an evidentiary hearing in the first instance. The first is whether AEDPA deference applies in this case to the state courts' determination at the first step of the inquiry required by *Batson*. We answer this question in the negative, which raises a second question: did Johnson and Thompson, on the basis of the state record, make the requisite prima facie showing of discrimination? We answer that question in the affirmative.

A

Under AEDPA, no federal court may grant a writ of habeas corpus unless the state courts adjudicated the petitioner's claim in a manner that “was contrary to, or involved an unrea-

sonable application of, clearly established Federal law, as determined by the Supreme Court of the United States.” 28 U.S.C. § 2254(d)(1). “When a state court’s adjudication of a claim is dependent on an antecedent unreasonable application of federal law,” however, “the requirement set forth in § 2254(d)(1) is satisfied. A federal court must then resolve the claim without the deference AEDPA otherwise requires.” *Panetti v. Quarterman*, 551 U.S. 930, 953 (2007). The question here is whether the state courts’ adjudication of Johnson and Thompson’s *Batson* claim was “dependent on an antecedent unreasonable application of federal law,” *id.*—namely, whether the state courts applied the proper standard in determining whether Johnson and Thompson made a *prima facie* showing of racial discrimination. In answering that question, “[w]e review the state court’s last reasoned decision,” *Crittenden v. Ayers*, 624 F.3d 943, 950 (9th Cir. 2010), which was in this case a decision made by the California Court of Appeal.

[1] At the first step of the *Batson* inquiry, a defendant need only “raise an inference that the prosecutor . . . exclude[d] the veniremen from the petit jury on account of their race.” 476 U.S. at 96 (emphasis added). The Court of Appeal recited the correct standard: “A party establishes a *prima facie* showing of invidious group bias when there is a reasonable inference from the circumstances as a whole that this was the basis for the peremptory challenge.” But the case that the court cited for that proposition was *People v. Box*, 5 P.3d 130 (Cal. 2000), *overruled on other grounds by People v. Martinez*, 224 P.3d 877 (Cal. 2010), which stated that “in California, a ‘strong likelihood’ means a ‘reasonable inference.’” *Id.* at 154 n.7. The U.S. Supreme Court squarely rejected that doctrine of California law as contrary to *Batson*. See *Johnson v. California*, 545 U.S. 162 (2005), *rev’g People v. Johnson*, 71 P.3d 270, 277 (Cal. 2003) (“We reiterate what we . . . stated in *Box*: . . . ‘strong likelihood’ and ‘reasonable inference’ state the same standard.”). In *Johnson*, the Court quoted with approval the California Court of Appeal’s statement that to

equate a “strong likelihood” and a “reasonable inference” is “as novel a proposition as the idea that ‘clear and convincing evidence’ has always meant a ‘preponderance of the evidence.’ ” 545 U.S. at 166 n.2. A state court that equates a correct standard with an incorrect standard cannot be applying the correct standard in the manner required by law.

Moreover, the Court of Appeal’s reasoning here leaves little doubt that in equating a “strong likelihood” with a “reasonable inference,” it was improperly heightening the latter standard rather than diminishing the former. The version of the “reasonable inference” standard that the Court of Appeal applied was the one rejected as unlawful in *Johnson*, not the one recognized by federal law. The strongest evidence of the court’s error is its statement that “[w]hen a trial court denies a motion to contest the basis of a peremptory challenge because there is no prima facie showing,” the appellate court must affirm so long as “there are grounds upon which a prosecutor could reasonably have premised a challenge.” As we explained in *Williams v. Runnels*, 432 F.3d 1102 (9th Cir. 2006), while “other relevant circumstances” can “rebut an inference of discriminatory purpose based on statistical disparity,” these “‘other relevant circumstances’ must do more than indicate that the record would support race-neutral reasons for the questioned challenges.” *Id.* at 1107-08. Contrary to the Court of Appeal’s reasoning, the existence of “grounds upon which a prosecutor could reasonably have premised a challenge,” does not suffice to defeat an inference of racial bias at the first step of the *Batson* framework.

[2] The only remaining question is whether the federal law that the Court of Appeal failed to apply reasonably was clearly established by the Supreme Court at the time of the Court of Appeal’s decision, as AEDPA requires in order for the state court’s error to be a basis for declining deference. The state argues that because *Johnson* was decided in 2005, three years after the state court of appeal decided this case, “there was no United States Supreme Court decision” reject-

ing as erroneous California's "longstanding holdings" that a strong likelihood and a reasonable inference had the identical meaning. Br. at 35-36. But in *Williams*, we rejected precisely the same argument: there, we held that we did not owe deference to state court decisions issued prior to *Johnson*, and using the "strong likelihood" standard, because "the Supreme Court clearly indicates in *Johnson* that it is clarifying *Batson*, not making new law." 432 F.3d at 1105 n.5; see *Johnson*, 545 U.S. at 169 (observing that "*Batson* . . . on its own terms provides no support for California's rule"). *Williams* explains why the federal law that the California Court of Appeal applied unreasonably here is *Batson* itself, not just its restatement in *Johnson*. We are bound by, and we agree with, *Williams*'s holding that "where the state court used the 'strong likelihood' standard for reviewing a *Batson* claim, the state court's findings are not entitled to deference." *Id.* at 1105 (citing *Paulino v. Castro*, 371 F.3d 1083, 1090 (9th Cir. 2004)); see also *Fernandez v. Roe*, 286 F.3d 1073, 1077 (9th Cir. 2002); *Cooperwood v. Cambra*, 245 F.3d 1042, 1046 (9th Cir. 2001).

In an appeal of the denial of a habeas petition without AEDPA deference, "we review de novo questions of law and mixed questions of law and fact. Factual findings and credibility determinations that were not made by the [state] trial court but were made by the district court after an evidentiary hearing are reviewed for clear error." *Crittenden*, 624 F.3d at 954 (citations omitted).¹ We review de novo the question whether the district judge deprived Johnson and Thompson of

¹The evidentiary hearing in this case was not of the sort barred by *Cullen v. Pinholster*, 131 S. Ct. 1388 (2011), which held that "review under [28 U.S.C.] § 2254(d)(1) is limited to the record that was before the state court that adjudicated the claim on the merits." *Id.* at 1398. If *Pinholster* were applicable, it would have been improper for the district or magistrate judge to take new evidence in determining whether the state courts' handling of the petitioners' *Batson* claim "was contrary to, or involved an unreasonable application of, clearly established Federal law," 28 U.S.C. § 2254(d)(1). But because the magistrate judge properly determined that the California Court of Appeal's decision did *not* qualify for deference under that provision, it was both lawful and necessary—just as in *Crittenden*—to conduct an evidentiary hearing in order to resolve the *Batson* claim by addressing the issues that the state court (as a result of its erroneous analysis) failed to reach.

the process that they were constitutionally due when it rejected the magistrate judge's credibility determination without conducting a new evidentiary hearing. *Ridgway*, 300 F.3d at 1155.

B

Having concluded that we owe no AEDPA deference to the state courts' determination that Johnson and Thompson failed to make a prima facie showing of racial discrimination, we must determine—de novo, *Crittenden*, 624 F.3d at 954 — whether the petitioners have shown that the evidence relating to the voir dire process at their trial, including all relevant circumstances, raises an inference of racial bias in the prosecution's exercise of its peremptory strikes.

[3] *Batson* explained how a defendant may make such a case:

[A] defendant may establish a prima facie case of purposeful discrimination in selection of the petit jury solely on evidence concerning the prosecutor's exercise of peremptory challenges at the defendant's trial. To establish such a case, the defendant first must show that he is a member of a cognizable racial group, and that the prosecutor has exercised peremptory challenges to remove from the venire members of the defendant's race. Second, the defendant is entitled to rely on the fact, as to which there can be no dispute, that peremptory challenges constitute a jury selection practice that permits "those to discriminate who are of a mind to discriminate." Finally, the defendant must show that these facts and any other relevant circumstances *raise an inference* that the prosecutor used that practice to exclude the veniremen from the petit jury on account of their race.

476 U.S. at 96 (citations omitted). We have recognized that “a defendant can make a prima facie showing based on statistical disparities alone.” *Paulino*, 371 F.3d at 1091.

[4] The fact that “three of the prosecution’s peremptory challenges were exercised against the only three African-Americans in the jury pool,” is enough to establish a prima facie case of racial discrimination. In multiple cases, we have held that a prima facie showing of racial discrimination had been made where prosecutors had stricken a lesser proportion of the racial minorities in a venire pool. *See, e.g., Paulino*, 371 F.3d at 1091 (finding a prima facie showing where “the prosecution had struck five out of six possible black jurors”); *Fernandez*, 286 F.3d at 1078 (finding a prima facie showing where “[t]he prosecutor [had] struck four out of seven . . . Hispanics” and “the only two prospective African-American jurors”); *Turner v. Marshall*, 63 F.3d 807, 812 (9th Cir.1995), *overruled on other grounds by Tolbert v. Page*, 182 F.3d 677, 681 (9th Cir.1999) (en banc) (finding a prima facie showing where “the prosecutor had used peremptory challenges to exclude five African-Americans out of a possible nine African-American venirepersons”). As the Supreme Court observed in *Miller-El v. Cockrell*, 537 U.S. 322 (2003), in which the prosecutor had exercised peremptory strikes against ten out of the eleven black jurors not removed by strikes for cause or by agreement, “[h]appenstance is unlikely to produce this disparity.” *Id.* at 342. The same is true where the prosecutor used peremptory strikes to remove *all* of the black jurors from the venire pool.

[5] The Supreme Court has made clear that it “did not intend the first step” of the *Batson* inquiry “to be so onerous that a defendant would have to persuade the judge—on the basis of all the facts, some of which are impossible for the defendant to know with certainty—that the challenge was more likely than not the product of purposeful discrimination.” *Johnson*, 545 U.S. at 170. A defendant makes a prima facie showing at *Batson*’s first step merely by “producing evi-

dence sufficient to *permit* the trial judge to draw an *inference* that discrimination has occurred.” *Id.* (emphasis added). The prosecutor’s use of peremptory strikes to remove all of the black potential jurors in the venire pool for Johnson and Thompson’s trial clearly raised a reasonable inference of racial discrimination.

It is true that statistical disparity alone does not end the inquiry; *Batson* held that we must “consider *all* relevant circumstances.” 476 U.S. at 96 (emphasis added). As we noted earlier, however, such “‘relevant circumstances’ must do more than indicate that the record would support race-neutral reasons for the questioned challenges.” *Williams*, 432 F.3d at 1108. The state obviously misunderstands that principle in presenting, as “relevant circumstances,” the argument that “there were numerous legitimate race-neutral reasons for the prosecutor to excuse each of the . . . prospective jurors,” Br. at 37. The consideration that the state urges belongs at the later steps of the *Batson* inquiry, when the prosecutor is required to proffer race-neutral reasons for the strike and the court is required to determine whether those explanations are credible.² The existence of “legitimate race-neutral reasons” for a peremptory strike, *id.*, can rebut at *Batson*’s second and third steps the *prima facie* showing of racial discrimination that has been made at the first step. But it cannot negate the existence of a *prima facie* showing in the first instance, or else the Supreme Court’s repeated guidance about the minimal burden of such a showing would be rendered meaningless.

²*See Tolbert v. Page*, 182 F.3d 677, 680 (9th Cir. 1999) (en banc) (“First, the [*Batson*] movant must make a *prima facie* showing that the prosecution has engaged in the discriminatory use of a peremptory challenge . . . Second, if the trial court determines a *prima facie* case has been established, the burden shifts to the prosecution to articulate a race-neutral explanation for challenging the juror in question. Third, if the prosecution provides such an explanation, the trial court must then rule whether the movant has carried his or her burden of proving the existence of purposeful discrimination.”).

The state's other argument on this point is similarly incorrect. Because the magistrate and district judges "ultimately acknowledged the propriety of excusing the second and third prospective jurors," the state argues, we are "left with a statistical analysis in which the prosecutor used his seventh peremptory challenge to excuse a lone African American prospective juror." Br. at 37. But the state's argument again ignores the difference between step one and the later steps of the *Batson* framework. It is true that the magistrate judge, having found a prima facie case of racial discrimination at step one, concluded at step three that the prosecution had stricken two black jurors for genuine race-neutral reasons. Contrary to the state's understanding, however, that ultimate conclusion does not negate the existence of a prima facie case in the first place. The *Batson* framework is one of burden-shifting. The party that objects (here, the defendant) bears the burden at steps one and three; the other side (here, the state) bears the burden at step two. These steps must be taken in the proper sequence. That a defendant fails to meet his burden at step three does not mean that he failed to meet his burden at step one. The magistrate and district judges found that the petitioners did not meet their ultimate burden of showing that the prosecutor's race-neutral reasons for striking the two black jurors other than W.J. were pretextual, *notwithstanding* the prima facie showing that these jurors were stricken for illegitimate reasons. The state is mistaken in arguing that this ultimate conclusion as to two jurors negates the district court's finding of a prima facie case of racial discrimination as to all three black jurors.

[6] On de novo review, we agree with the magistrate and district judges that Johnson and Thompson did make a prima facie showing of racial discrimination at the first step of the *Batson* framework. It was therefore the duty of the magistrate judge to conduct an evidentiary hearing, in order to replicate on habeas review the inquiry that the state trial court should have conducted in the first place—requiring the prosecutor to assert race-neutral reasons for the strike (at *Batson* step two)

and determining (at *Batson* step three) whether the asserted reasons were in fact genuine rather than pretextual. Because the reasons that the prosecutor proffered for striking W.J. were race-neutral on their face, we proceed to consider the central question in this appeal: whether the district judge properly handled the inquiry required by *Batson*'s third step.

III

[7] Under 28 U.S.C. § 636(b)(1), when a district judge delegates to a magistrate judge the task of conducting an evidentiary hearing concerning a habeas petition, the district judge is to “make a de novo determination of those portions of the [magistrate judge’s] report or specified proposed findings or recommendations to which objection is made.” *Id.* § 636(b)(1)(C). In two cases concerning magistrate judge rulings on motions to suppress, however, we have held as a matter of constitutional due process “that a district court must conduct its own evidentiary hearing before rejecting a magistrate judge’s credibility findings.” *United States v. Ridgway*, 300 F.3d 1153, 1154 (9th Cir. 2002).³ We initially adopted this rule in *United States v. Bergera*, 512 F.2d 391 (9th Cir. 1975), explaining that a requirement for “the district court to rehear the evidence if it decides not to follow the recommendations of the magistrate insures that any decision on the facts will be the result of first-hand observation of witnesses and evidence.” *Id.* at 393. As we stated in *Bergera*, “[t]he law has long recognized the value of these more immediate impressions, and gives them a measure of protection from easy modifications made on the basis of dry records.” *Id.* at 393.

Ridgway reaffirmed this rule, and explained its constitutional foundation, in light of the Supreme Court’s decision in

³Magistrate judge rulings on motions to suppress, like those concerning habeas petitions, fall within the class of rulings authorized by 28 U.S.C. § 636(b)(1)(B), all of which the district judge reviews de novo under § 636(b)(1)(C). *See Ridgway*, 300 F.3d at 1154.

United States v. Raddatz, 447 U.S. 667 (1980). The Court held in *Raddatz* that a district judge could *accept* a magistrate judge's determination of credibility without holding a new evidentiary hearing, while expressing doubt as to whether a district judge could *reject* a magistrate judge's finding in these circumstances. The Court stated in a footnote that it found the latter prospect troubling: "[W]e assume it is unlikely that a district judge would reject a magistrate's proposed findings on credibility when those findings are dispositive and substitute the judge's own appraisal; to do so without seeing and hearing the witness or witnesses whose credibility is in question could well give rise to serious questions which we do not reach." *Id.* at 681 n.7.

[8] Although we have not yet explicitly extended this doctrine beyond rulings on motions to suppress, its rationale clearly applies to *Batson* motions by criminal defendants. As the Supreme Court has explained, it is essential that judges who rule at *Batson*'s third step have the opportunity to witness the prosecutor's testimony in person: "In the typical peremptory challenge inquiry, the decisive question will be whether counsel's race-neutral explanation for a peremptory challenge should be believed. There will seldom be much evidence bearing on that issue, and the best evidence often will be the demeanor of the attorney who exercises the challenge." *Hernandez v. New York*, 500 U.S. 352, 365 (1991); *see also Gomez v. United States*, 490 U.S. 858, 874-75 (1989) ("To detect prejudices [during voir dire], . . . [t]he court . . . must scrutinize not only spoken words but also gestures and attitudes of all participants to ensure the jury's impartiality."); *United States v. You*, 382 F.3d 958, 968 (9th Cir. 2004) ("A trial court's findings on purposeful discrimination rest largely on credibility. Courts measure credibility 'by, among other factors, the prosecutor's demeanor . . .'" (citation omitted)). "There can be no doubt," we have held, "that seeing a witness testify live assists the finder of fact in evaluating the witness's credibility. . . . Live testimony enables the finder of fact to see the witness's physical reactions to questions, to assess the wit-

ness's demeanor, and to hear the tone of the witness's voice—matters that cannot be gleaned from a written transcript.” *United States v. Mejia*, 69 F.3d 309, 315 (9th Cir. 1995). A district judge who rejects a magistrate judge's finding as to the credibility of a prosecutor's explanation for a peremptory strike, without seeing the prosecutor testify in person, is just as hampered by the deficiencies of a cold record as one who rejects a magistrate judge's finding as to the credibility of testimony in a suppression hearing.

Indeed, the Supreme Court has suggested in two cases that the considerations discussed in the *Raddatz* footnote extend to the context of voir dire. First, in holding that magistrate judges could not preside over voir dire in a felony trial without the defendant's consent, the Court commented in a footnote:

Like motions to suppress evidence, petitions for writs of habeas corpus, and other dispositive matters entailing evidentiary hearings, jury selection requires the adjudicator to observe witnesses, make credibility determinations, and weigh contradictory evidence. Clearly it is more difficult to review the correctness of a magistrate's decisions on these matters than on pretrial matters, such as discovery motions, decided solely by reference to documents.

Gomez v. United States, 490 U.S. 858, 874 n.27 (1989) (citation omitted). Then, in its subsequent and related holding that magistrate judges do have the power to supervise felony voir dire *with* the defendant's consent, the Court acknowledged that “*de novo* review by the district court” might in certain cases “provide an inadequate substitute for the Article III judge's actual supervision of the voir dire.” *Peretz v. United States*, 501 U.S. 923, 939 (1991). But “the same,” it said, was “true of a magistrate's determination in a suppression hearing, which often turns on the credibility of witnesses,” and which

Raddatz expressly authorized. *Id.*⁴ In other words, the Court in these cases understood the constitutional problem in the voir dire and suppression contexts to be the same: because a determination in these matters generally relies on the ability to observe a witness, it is difficult—and constitutionally troubling—for a district judge to disagree with the determination reached by a magistrate judge without first hearing the relevant testimony in person.

[9] Taking the Supreme Court’s various hints, the First, Second, Third, Fifth, and Eleventh Circuits have all held that a district judge may not reject the credibility finding of a magistrate judge without holding a new evidentiary hearing. *See Louis v. Blackburn*, 630 F.2d 1105, 1109 (5th Cir. 1980) (“[I]n a situation involving the constitutional rights of a criminal defendant, we hold that the district judge should not enter an order inconsistent with the credibility choices made by the magistrate without personally hearing the live testimony of the witnesses whose testimony is determinative.” (footnote omitted)); *Hill v. Beyer*, 62 F.3d 474, 482 (3d Cir. 1995) (“A district court may not reject a finding of fact by a magistrate judge without an evidentiary hearing, where the finding is based on the credibility of a witness testifying before the magistrate judge and the finding is dispositive of an application for post-conviction relief involving the constitutional rights of a criminal defendant.”); *Cullen v. United States*, 194 F.3d 401, 407 (2d Cir. 1999) (“[I]t appears that a district judge should normally not reject a proposed finding of a magistrate judge that rests on a credibility finding without having the witness testify before the judge.”); *United States v. Hernandez-Rodriguez*, 443 F.3d 138, 148 (1st Cir. 2006) (“[W]e join our sister circuits when we find that, absent special circumstances, a district judge may not reject the credibility determination of a magistrate judge without first hearing

⁴As in the *Raddatz* footnote, the Court in *Peretz* “presume[d] . . . that district judges [would] handle such cases properly if and when they [were to] arise.” *Id.*

the testimony that was the basis for that determination.”); *United States v. Cofield*, 272 F.3d 1303, 1306 (11th Cir. 2001) (“[G]enerally a district court must rehear the disputed testimony before rejecting a magistrate judge’s credibility determinations.”).⁵ We agree with these circuits that the rationale of *Ridgway* and the *Raddatz* footnote applies generally to determinations affecting the rights of a criminal defendant and involving a credibility finding. A district court may not in such instances reject a magistrate judge’s proposed credibility determination without hearing and seeing the testimony of the relevant witnesses.

The state’s only response to Johnson and Thompson’s arguments concerning *Ridgway* is to assert, in a single footnote, that “*Ridgway* is wholly inapplicable here because the Magistrate Judge’s factual findings regarding the prosecutor were purely based upon his crabbed comparative analysis and not upon any observations of the prosecutor’s demeanor while testifying.” Br. at 18 n.8. We explicitly rejected this argument, however, in *Ridgway* itself. There, the district judge had asserted—much as the state does here—that the relevant witness’s credibility “could be assessed by reviewing the cold record, without personally observing the witness, because ‘the magistrate judge ha[d] founded his credibility determination upon supposed discrepancies, not the witness’s demeanor or any other attribute which is unavailable in the paper record.’ ” 300 F.3d at 1155. We disagreed, on the basis that “[t]he broad rule announced in *Bergera* contains no exceptions,” 300 F.3d at 1157, and we believe our holding in *Bergera* applies with equal force in the *Batson* context.

⁵The Sixth, Seventh, Eighth, and Tenth Circuits have found it unnecessary to reach the question in cases before them. See *United States v. Bailey*, 302 F.3d 652, 657 n.5 (6th Cir. 2002); *United States v. Ornelas-Ledesma*, 16 F.3d 714, 720 (7th Cir. 1994), *vacated on other grounds*, *Ornelas v. United States*, 517 U.S. 690 (1996); *United States v. Black Bear*, 422 F.3d 658, 662 n.1 (8th Cir. 2005); *United States v. Orrego-Fernandez*, 78 F.3d 1497, 1502 (10th Cir. 1996). No circuits appear to have rejected the rule in question.

[10] Even aside from its conflict with our precedent, the state's argument is erroneous because a district judge's review of a magistrate judge's credibility finding is in no way limited to the specific reasons offered by the magistrate judge. A magistrate judge might choose to explain his adverse credibility finding on the basis of the paper record, even though he also considers a witness's demeanor to be suspect. A credibility determination—particularly in a *Batson* challenge—ordinarily involves the fact-finder's assessment of the witness's demeanor as well as his review of the record. *See Hernandez*, 500 U.S. at 365; *see also Mejia*, 69 F.3d at 315. A district judge who disagrees with the magistrate judge's written analysis of the record might nonetheless, if he took the time to observe the witness in person, agree with the magistrate judge's unwritten assessment of the witness's demeanor and affirm the magistrate judge's overall credibility determination on that basis. "If the district judge doubts the credibility determination of the magistrate, only by hearing the testimony himself does he have an adequate basis on which to base his decision." *Louis*, 630 F.2d at 1110.

[11] We therefore hold that the district judge deprived Johnson and Thompson of the process that they were constitutionally due, when he rejected the magistrate judge's proposed finding as to the prosecutor's lack of credibility without observing the prosecutor's testimony in person. "The guarantees of due process call for a 'hearing appropriate to the nature of the case.'" *Raddatz*, 447 U.S. at 677 (quoting *Mullane v. Central Hanover Bank & Trust Co.*, 339 U.S. 306, 313 (1950)). The nature of this case, like every case that reaches the third step of the *Batson* analysis, demands that the ultimate trier of fact hear testimony in person: "the decisive question" is "whether counsel's race-neutral explanation for a peremptory challenge should be believed," and "the best evidence" regarding that question "will be the demeanor of the

attorney who exercises the challenge.” *Hernandez*, 500 U.S. at 365.⁶

[12] Johnson and Thompson were constitutionally entitled to have the district judge observe the prosecutor’s demeanor before rejecting their *Batson* claim. Because the petitioners’ interest in the vindication of their rights is immense, because the administrative burden of an additional hearing is relatively minor, and because a credibility determination based on a cold record is substantially more likely to be in error than one based on an in-person evaluation of a witness, the district judge deprived Johnson and Thompson of due process when he declined to afford them a new evidentiary hearing. *See Mathews v. Eldridge*, 424 U.S. 319, 335 (1976) (enumerating the factors to be weighed in a constitutional due process analysis); *Louis*, 630 F.2d at 1110 (applying the *Mathews* factors in holding that a district judge may not reject a magistrate judge’s credibility determination without holding a new evidentiary hearing).

IV

[13] Johnson and Thompson contend that the proper remedy is for us to look through the district judge’s order to review for clear error the magistrate judge’s credibility determination. We disagree. *See Cullen*, 194 F.3d at 407 (holding

⁶The prosecutor’s demeanor might be particularly useful evidence of his credibility here, where his testimony included the “recit[ation of] a laundry list of reasons,” for the peremptory strike of W.J.—a few of which verge on “implausible or fantastic,” *Purkett v. Elem*, 514 U.S. 765, 768 (1998); many of which were not subjects of inquiry by the prosecutor during voir dire, *cf. Ali v. Hickman*, 584 F.3d 1174, 1192 (9th Cir. 2009); and some of which are unlikely to hold up under a comparative juror analysis, *cf. Green v. Lamarque*, 532 F.3d 1028, 1030 n.3 (9th Cir. 2008). Because “[a] *Batson* challenge does not call for a mere exercise in thinking up any rational basis,” *Miller-El v. Dretke*, 545 U.S. 231, 252 (2005), the district judge must “evaluate *meaningfully* the persuasiveness of the prosecutor’s [race]-neutral explanations,” *United States v. Alanis*, 335 F.3d 965, 969 (9th Cir. 2003).

that simply to review the magistrate judge's determination "would elevate the recommended ruling of the Magistrate Judge to a final ruling and undermine section 636(b)(1)'s requirement of a de novo determination by the District Court"). As in *Ridgway*, we vacate the district judge's order and remand for the district judge either to adopt the magistrate judge's credibility determination or to conduct a new evidentiary hearing.⁷ We retain jurisdiction over any appeal from the judgment on remand.

VACATED and REMANDED.

⁷The First, Second, and Third Circuits have in similar cases ordered that a different district judge conduct the required hearing on remand. See *Cullen*, 194 F.3d at 408; *Hernandez-Rodriguez*, 443 F.3d at 148; *Boyd v. Waymart*, 579 F.3d 330, 333 (3d Cir. 2009) (en banc) (per curiam); see also *id.* at 339 & n.10 (Scirica, C.J., concurring) (citing *Cullen*, 194 F.3d at 408). Because Johnson and Thompson have not asked that this case be reassigned, however, and the parties have not briefed the issue, we do not decide whether that remedy would be appropriate here. Nor do we express any view as to whether reassignment is the proper remedy in similar cases when requested by the appellant.