STATEMENT BY ROGER CLEGG,
VICE PRESIDENT AND GENERAL COUNSEL OF
THE CENTER FOR EQUAL OPPORTUNITY,
BEFORE THE U.S. COMMISSION ON CIVIL RIGHTS,
RE RACIAL AND ETHNIC DATA COLLECTION
BY GOVERNMENT AGENCIES
MAY 17, 2002

Thank you, Madame Chairman, for the opportunity to testify before the Commission today about the collection of racial and ethnic data by government agencies, and in particular its effect on civil-rights enforcement.

My name is Roger Clegg, and I am vice president and general counsel of the Center for Equal Opportunity, a nonprofit, nonpartisan, Section 501 (c)(3) research and educational organization based in Sterling, Virginia. Our president is Linda Chavez, who used to be the Commission's director. At CEO, I focus on civil rights issues. I also served for over ten years at the U.S. Department of Justice, including four years in the civil rights division, during the Reagan and earlier Bush administrations.

Let me say at the outset that my testimony will not categorically defend or condemn racial and ethnic data collection. For example, sometimes the collection of such data can be useful in civil-rights enforcement, but sometimes it undermines civil-rights enforcement by facilitating discrimination. More broadly, sometimes such data can be put to good use, sometimes it can be abused, and some methods of data collection are more problematic than others.

Thus, it would be misleading to consider only the good uses to which data collection can be put without also considering the problems with such collection. Whether the positives outweigh the negatives depends on which government agency we're talking about, what the agency is likely to do with the data, and how it intends to go about collecting it. Thus, I support the Racial Privacy Initiative that the American Civil Rights Coalition is trying to place before the voters in California, which would ban some data collection, but would note that the RPI correctly creates some exceptions to an across-the-board ban.

Those who want the government to put people into what Ward Connerly has called the "silly little boxes" on written forms should bear the burden of proof in showing that there are very good reasons for needing this information, and that the likelihood of misuse is small. I will discuss the problems with data collection first, and then some of the possible benefits.

Harm from Racial and Ethnic Data Collection

There are two basic kinds of problems that result from the collection of these data. The first arise from the very process of collecting the data, and the second from the misuse of the data once they are collected.

The process of collecting the data, in turn, also gives rise to different subproblems. People are encouraged to think of themselves as having a particular racial identity, and the government is encouraged to require that they have such an identity. It is difficult to solve these two problems when racial and ethnic data are being collected, and indeed it may be impossible to solve them simultaneously. Requiring the government to do the classifying rather than the individual is likely to lead to errors, and of course it would be offensive to train officials in how to identify people by racial or ethnic "characteristics." On the other hand, it is also intrusive for the government to demand that people identify themselves racially and, no matter how "voluntary" you make such identification, when the government is asking for it, pressure is being brought to bear.

The problems with requiring people to identify themselves racially and ethnically become more problematic in a multiracial, multiethnic society with much intermarriage. If we lived in a world in which each person had to be either completely white or completely black, then you might say that the government's requirement that people check a box for one or the other really isn't forcing an identity on them that they don't already have. But of course that is not the case in America today. We have lots of groups within blacks and whites and besides blacks and whites, and we have more and more individuals who are racially and ethnically blended. In California, for instance, interracial births are the third highest category of births, behind white and Latino.

It is offensive for the government to ask these children about their racial and ethnic background and demand that they choose a particular identification. It is really none of the government's business, just as religion and sexual orientation are generally none of the government's business. It is interesting that France has long honored this principle, and "collects no systematic data on race or ethnicity." Erik Bleich, "The French Model: Color-Blind Integration," in *Color Lines: Affirmative Action, Immigration, and Civil Rights Options for America* (John David Skrentny, editor, University of Chicago Press 2001), at 286.

Insisting that people embrace a racial identity is bad for civil-rights progress and, therefore, bad for civil-rights enforcement. Discrimination is more likely to occur in a society in which people have strong racial identities and an us-them mentality. I understand that no American is literally color-blind and that many people feel like they have their racial identities given to them by society and see nothing wrong with the government recognizing that reality. But we ought to be doing what we can to move away from such racial identifications, and it is certainly a bad idea for the government to be encouraging this kind of categorization. Getting the government out of this business sends a strong, positive message that we are all Americans and that skin color and ancestry don't matter here.

Another problem that arises from the very process of data collection is that the government may be encouraged to engage in racial discrimination if it wants a particular set of racial results. For instance, suppose that a police department is required to keep

track of the race and ethnicity of the people pulled over for traffic violations. Such a requirement is frequently proposed as a means of combating racial profiling by the police. The trouble is that policemen know that they may get into trouble if a disproportionate number of their stops involve, say, African Americans—even if the disproportion is not a result of any discrimination on the part of the police. A policeman, therefore, may be reluctant to pull over a black motorist who is speeding if he has already pulled over a couple of other black motorists that evening; conversely, he may decide to pull over a nonblack motorist in a situation when he would have let a black motorist go. I believe that the District of Columbia police chief—who happens to be black—recently made this very point in a radio interview.

Or suppose you have an agency that is told to keep track of the race and ethnicity of the people it hires, in order to ensure that there is no discrimination. Suppose managers are also told that their performance evaluation will depend, in part, on their commitment to equal employment opportunity. The trouble is that a manager who has already hired several Latinos may be reluctant to hire another person of that background if she thinks that the resulting "disproportion" will call into question her commitment to nondiscrimination. Of course, the problem is aggravated if the manager's evaluation is not rooted in whether she is committed to nondiscrimination, but whether she is committed to "diversity." Then she is basically being told that she needs to take race and ethnicity into account in her hiring.

In such situations, you can see how the very process of collecting racial and ethnic data for ostensibly civil-rights enforcement reasons actually ends up encouraging the violation of civil-rights laws and principles.

The second category of problems arises when a government agency deliberately uses the data it has already collected in order to discriminate. When a college admissions office collects such data, the data are almost certainly going to be used to discriminate in favor of some individuals on the basis of skin color or ancestry and against others. There is really no valid use to which the admissions office can put this information. And I have already discussed how personnel data are frequently used to ensure that groups are proportionately represented.

Obviously, in these situations the collection of data actually undermines civil-rights enforcement, because the data facilitate the violation of the civil-rights laws.

Legitimate Uses of Racial and Ethnic Data

On the other hand, there are situations in which the collection of racial and ethnic data is legitimate and useful. A good example is when a prison includes a physical description of an inmate in his file; this information would obviously be useful if, say, the prisoner escaped and the police needed to identify him. Another example is in some kinds of scientific and academic research. A study of how well patients respond to a new drug may want to keep track of race and ethnicity, in case that has an effect on the drug's effectiveness. *See* Sally Satel, "I Am a Racially Profiling Doctor," *New York Times*, May

5, 2002. Social scientists and historians also use and analyze racial and ethnic data, and I am prepared to believe that some of it, at least, has to be collected by the government rather than by private parties. These data can provide important insights, not least of which is documenting the extent of progress being enjoyed by racial and ethnic minority groups and thereby counteracting the paranoia of many doomsayers. University of California professor John McWhorter and Commissioner Thernstrom have made this point.

Note that these legitimate uses of racial and ethnic data are not directly related to civil-rights enforcement; nonetheless, they are important.

The other major and legitimate use of racial and ethnic data is in determining whether racial or ethnic discrimination has occurred. For instance, my organization—the Center for Equal Opportunity—collects admissions data from public colleges and universities under state freedom-of-information laws, and then subjects the data to a multiple logistic regression analysis to determine if the schools are engaged in racial or ethnic discrimination (our studies, which have found a great deal of discrimination all over the country, in undergraduate, medical school, and law school admissions, can be found on our website, www.ceousa.org).

Obviously, we would be unable to conduct these studies unless the schools kept racial and ethnic information. On the other hand, the schools would probably find it more difficult to engage in discrimination in the first place if they lack these data. As I said earlier, on balance I support the Racial Privacy Initiative, even though in this one area it would make my job harder, because I think that the RPI would diminish the amount of racial and ethnic discrimination that occurs.

Finally, racial and ethnic data can also be used in the course of a lawsuit. There are three important caveats here, though. First, many civil-rights lawsuits do not hinge on the use of statistical data at all. Second, it is also possible to collect the information after a lawsuit has been filed, through the discovery process. And third, statistics can be used and abused in lawsuits. I have been very critical of "disparate impact" lawsuits—that is, lawsuits that do not allege disparate treatment because of race, but simply challenge the use of selection criteria that have a disproportionate "effect" on one group or another. *See* Roger Clegg, *Disparate Impact in the Private Sector: A Theory Going Haywire* (National Legal Center for the Public Interest 2001). In my view, making such lawsuits harder to bring would be a good thing, not a bad thing, because their inevitable—and, I believe, intended—result is to encourage the use of racial and ethnic quotas and to discourage the use of perfectly legitimate selection criteria. Thus, while the collection of racial and ethnic data may facilitate some civil-rights lawsuits, they do so at the expense of encouraging civil-rights violations. On balance, it is not clear to me that civil-rights enforcement through litigation is furthered by data collection.

I would point out that data are typically not collected about people's religion, yet there is no claim that the civil-rights laws' ban on religious discrimination has been undermined by this fact. Likewise, the Employment Non-Discrimination Act, which

would create a federal ban on employment discrimination on the basis of sexual orientation, would at the same time ban the collection of statistics by the EEOC with respect to sexual orientation. I oppose ENDA, but even those who support it think that it will work perfectly well without data collection. The point is that questions about religion and sexual orientation are viewed as none of the government's business, even if they might sometimes facilitate a lawsuit. A person's melanin content and his or her ancestors' countries of origin is, likewise, none of the government's business.

Conversely, while the collection of racial and ethnic data can help tell us if racial and ethnic imbalances are present, they do not tell us whether those imbalances are the result of discrimination. For that, we also have to have data on all the other variables that might account for whether a person is selected for a job, school admission slot, or whatever. Unless the employer or school is keeping all these other data, too, then not collecting racial and ethnic data isn't denying much of value to potential litigants.

I should also point out that, even if you believe that someone should collect these data, it doesn't follow that it ought to be collected more than once. If the federal EEOC collects employment data—as it does—then it would not seem to be necessary for other government actors to do so. Every time you force someone to put himself or herself into one of the silly little boxes, you are reinforcing that message that a person's racial and ethnic identity is very important—a bad message.

The Harm from Data Collection Generally Outweighs the Benefits

As I said at the beginning of my statement, I think that weighing the costs and benefits of racial and ethnic data collection has to be done on a case by case basis, as the Racial Privacy Initiative has done. I hope that the initiative becomes law in California, and that we will be able to see whether any problems result from it. Frankly, I doubt that there will be any problems. The doomsayers opposed Proposition 209 (ending state preferences based on race, ethnicity, and sex) and Proposition 227 (ending bilingual education), but they turned out to be quite wrong, and I think they will be wrong this time, too.

But, in any event, the approach deserves a try, and our most populous and diverse state is the best place to try it. If it turns out that there are problems with it, RPI provides a means of amendment. But if it turns out to work well, then it can be a valuable model for the rest of the country.

If you think that government agencies and those they regulate ought to use racial and ethnic quotas and ought to be prosecuted if they fail to meet them, and if you think that the government and individuals should be encouraged to embrace racial and ethnic identification, then there is every reason to support the collection of racial and ethnic data by government agencies and no reason to oppose it. I suspect that the most vehement opponents of the Racial Privacy Initiative fall into this category.

On the other hand, if you think that racial and ethnic identification by individuals and, especially, the government is something we ought to be moving beyond, and if you think that quotas are a bad idea and that this sort of discrimination has now become as or more widespread than other forms of discrimination, then getting the government out of the business of racial and ethnic data collection starts to make a lot of sense.

Thank you, again, for the opportunity to discuss this issue with the Commission. I would be happy to try to answer any questions that the Commissioners have.