Electronic defects in metal oxide photocatalysts

Ernest Pastor, Michael Sachs, Shababa Selim, James R. Durrant, Artem A. Bakulin, and Aron Walsh

Abstract | A deep understanding of defects is essential for the optimization of materials for solar energy conversion. This is particularly true for metal oxide photo(electro)catalysts, which typically feature high concentrations of charged point defects that are electronically active. In photovoltaic materials, except for selected dopants, defects are considered detrimental and should be eliminated to minimize charge recombination. However, photocatalysis is a more complex process in which defects can have an active role, such as in stabilizing charge separation and in mediating rate-limiting catalytic steps. In this Review, we examine the behaviour of electronic defects in metal oxides, paying special attention to the principles that underpin the formation and function of trapped charges in the form of polarons. We focus on how defects alter the electronic structure of metal oxides, statically or transiently upon illumination, and discuss the implications of such changes in light-driven catalytic reactions. Finally, we compare oxide defect chemistry with that of new photocatalysts based on carbon nitrides, polymers and metal halide perovskites.

Solid-state photo(electro)chemistry uses semiconductors to capture sunlight and drive useful chemical reactions, such as the production of H₂ from water or the reduction of CO₂ into compounds with a higher energy density such as ethanol. Great progress has been made since the pioneering work by Gerischer on semiconductor–liquid interfaces and the milestone report of water splitting in TiO₂ by Fujishima and Honda, making today’s semiconductor photo(electro)chemistry an increasingly viable technology. The advances have been such that current research is no longer focused on merely demonstrating catalysis but is aiming to push conversion efficiencies into an economically competitive range and develop working devices.

The biggest strides in the field have been driven by improvements of well-known metal oxides, most notably through nanostructuring (for example, ‘cauliflower’ α-Fe₂O₃) and the assembly of heterojunctions (such as WO₃:BiVO₄) and the discovery of more efficient compositions (for example, SrTiO₃:Al nanoparticles). Crucially, all of these advances have been underpinned by improvements in the control of the concentration and distribution of defects, which — even in small amounts — can drastically change the properties of a solid and the performance of a device.

Defects are responsible for various phenomena detrimental to the electronic properties of energy conversion materials. For example, the structural deformations induced by defects can hinder charge transport and the electronic deformations can act as recombination centres that compromise device performance (such as quenched luminescence in Co-doped ZnO). However, despite their name, defects can also have a desirable influence on the material properties. For example, chemical doping can turn an insulating oxide into an efficient photocatalyst by increasing carrier concentrations (for example, doping TiO₂ with Nb increases the electron concentration) concomitantly generating ionized dopants as point defects in the structure. Similarly, control of electronic states associated with defects can help to tune and extend charge-carrier lifetimes.

Over recent years, there has been increasing awareness of the importance of polarons — namely charges that are trapped through lattice deformations even in the absence of a physical point defect — in the operation of photo(electro)catalytic devices. Understanding the role of defects and of localized polaronic charge carriers is helpful for improving functional solids for a broad range of applications, but is imperative in the case of heterogeneous photocatalysts.

In such systems, light triggers the reorganization of chemical bonds at a solid–liquid (or sometimes solid–gas) interface, requiring the participation of surface dangling bonds (that is, surface defects) for operation. Even more importantly, defect formation is in itself a chemical reaction that depends strongly on the chemical nature of the defects and their concentration.
of the solid. Complications emerge when the materials properties required to favour a desired chemical transformation are not notably different from those that favour defect formation.

In this Review, we discuss the impact that electronic defects have on the function of photo(electro)catalytic processes. We use the term ‘photocatalysis’ to refer to both photocatalytic and photoelectrocatalytic processes. Although the operation principles are different in these two cases (Box 1), herein we emphasize the common role that defects have in such technologies. By electronic defects, we refer to changes in the electronic structure induced by point defects (such as oxygen vacancies) and/or polaronic states associated with charge carriers. Although oxide photocatalysis is an extremely rich field in which promising materials are rapidly being discovered\textsuperscript{4,16–26} and new reactions are being catalysed (such as pollutant degradation and the reduction of CO\textsubscript{2} in both the liquid and gas phases\textsuperscript{11–31}, we refer primarily on established cases of photocatalytic oxides used for water splitting. In particular, we use α-Fe\textsubscript{2}O\textsubscript{3} as examples to provide a framework within which to understand how defect control can help increase photocatalytic yields in oxide-based systems.

We start by reviewing the defect chemistry in metal oxides, focusing on the formation of structural imperfections in the form of vacancies and polarons, and show how such processes alter the fundamental electronic structure of the catalysts. We then outline the key processes that light-driven catalysts must overcome to realize efficient photocatalysis, from the initial equilibration with the liquid phase to product formation. Next, we discuss how chemical imperfections affect such processes and emphasize how sub-bandgap energy levels affect the optoelectronic properties of the metal oxide and the kinetics and energetics of the reactive excited state. Finally, we compare the defect chemistry in oxides with that of emerging photocatalysts based on carbon nitrides, conjugated polymers and metal halide perovskites, and consider how the lessons learned from these systems can be applied to metal oxide systems.

**Point defects in metal oxides**

The formation of defects in solids is unavoidable and results in the loss of the translational symmetry of the crystallographic unit cell. Symmetry breaking can happen across multiple dimensions, giving rise to 3D volumetric defects (such as pores), 2D planar defects (such as grain boundaries), 1D linear defects (dislocations) and 0D point defects (such as vacancies)\textsuperscript{37,48}. All types of defect can affect the electronic and catalytic properties of metal oxides\textsuperscript{49–51}, but herein we focus on point defects as they are the most intrinsic and have a major role in the optimization of catalytic performance. For a comprehensive overview of the classification of point defects in crystals and associated characterization techniques, we refer readers to the seminal work of Kröger\textsuperscript{52}. We first provide an overview of the electronic structure of metal oxides and use it as a framework to discuss the chemical principles that underpin the formation of imperfections (both in the dark and upon illumination).

**Electronic structure and bandgaps**

The electronic structure and the magnitude of the bandgap of metal oxides is influenced by the crystal symmetry, the ionicity of the metal–oxygen bond and the degree of orbital mixing or hybridization\textsuperscript{53,54}. In general, ionic oxides with a large energy difference between the metal and oxygen atomic orbitals tend to exhibit large bandgaps outside the range of solar energy conversion\textsuperscript{15}. This is the case for ZnO (3.4 eV), Ga\textsubscript{2}O\textsubscript{3} (4.8 eV) and Al\textsubscript{2}O\textsubscript{3} (8.8 eV), which have electronic band structures formed by mixing of O 2p and metal s orbitals. By contrast, transition metal oxides with partially filled valence d orbitals, such as α-Fe\textsubscript{2}O\textsubscript{3} (2.2 eV) sustain additional bonding interactions that involve the metal d and/or s orbitals, which lower the bandgap. Often such changes are sufficient to absorb photons in the visible part of the solar spectrum, which makes these metal oxides prime candidates for photocatalysis.

As a first approximation, the electronic structure of most transition metal oxides can be explained from a crystal field perspective in which the oxygen anions are arranged in different geometries around a metal cation. Electrons that occupy the d orbitals of the metal centre experience different degrees of repulsion depending on the orbital orientation with respect to the surrounding oxygen anions. In an octahedral geometry (O\textsubscript{h}), metal atomic orbitals with t\textsubscript{2g} symmetry (d\textsubscript{xy}, d\textsubscript{yz}, d\textsubscript{zx}) are stabilized, while atomic orbitals with e\textsubscript{g} symmetry (d\textsubscript{z2}, d\textsubscript{x2−y2}) are destabilized (Fig. 1a). The removal of an oxygen atom from the octahedron, such as due to an oxygen vacancy or dangling bond at the surface, leads to a square pyramidal geometry (C\textsubscript{4v}) with a change in the splitting and stabilization of the t\textsubscript{2g} and e\textsubscript{g} orbitals\textsuperscript{43–47} (Fig. 1a).

Although crystal field theory provides an initial estimate of the electronic structure, this model does not describe covalent interactions. This extension is achieved with ligand field theory, which considers the overlap and mixing of oxygen 2p and metal 3d wavefunctions. Within this molecular orbital perspective, in an octahedral ligand field, metal atomic orbitals with t\textsubscript{2g} symmetry have weak overlap with ligand O 2p orbitals and form π-bonding and π*-antibonding molecular orbitals. By contrast, metal atomic orbitals with e\textsubscript{g} symmetry strongly interact with oxygen orbitals to form σ-bonding and σ*-antibonding molecular orbitals. FIGURE 1b shows a molecular orbital diagram for the building blocks of a transition metal oxide. The nature of the frontier orbitals will determine the redox chemistry and tendency for defect formation.

**Frontier orbitals**

Although it can be straightforward to define a small set of orbitals as the frontier orbitals (highest occupied molecular orbital (HOMO) and lowest unoccupied molecular orbital (LUMO)) in a molecule that are responsible for the chemical reactivity, in a macroscopic crystal, the picture is complicated by crystal momentum and the associated band dispersion in k-space. However, we can
**Box 1 | Solid-state photocatalytic approaches**

Solid-state photocatalysis is generally achieved using particulate systems in suspension and films or through photoelectrochemical (PEC) cells (see figure). In a particulate system, a single particle drives a complete redox reaction (oxidation and reduction), often with the aid of a co-catalyst. In such systems, catalytic particles are in close proximity, limiting losses from solution resistance and mass transport but also complicating the collection and separation of products. Traditionally, suspension photocatalysis has been performed in ‘baggie systems’ (see the figure, panel a), but this approach is also being implemented in the form of photocatalytic panels in which the particles are immobilized on support panels. In contrast to suspension systems, in PEC cells, the two reactions occur at separate electrodes — oxidation at the (photo)anode and reduction at the (photo)cathode — thereby facilitating product separation and collection (see the figure, panel b). There are multiple PEC cell configurations for, for example, the electrodes can be joined together to form a compact device or they can be physically separated and connected by wires.

![Image](https://example.com/image.png)

Equally, some systems employ semiconductors as both the photocathode and photoanode, whereas other systems employ just one photoelectrode and a second metallic electrode. The latter case requires the semiconductor to generate enough photovoltage to drive the full reaction (oxidation and reduction) or relies on the assistance of an external voltage source, such as a solar cell. It is important to note that, although the eventual goal is to design a complete functional system, the oxidation and reduction reactions are usually studied separately in order to optimize the active materials. This optimization is achieved by using sacrificial donors in suspensions or by voltage-assisted PEC cells that involve the desired photoelectrode and a metallic counter electrode (see the figure, panel c). The alignment of the conduction band (CB) and valence band (VB) with the redox potential (E) of the species in solution (inset) will dictate whether a reaction is possible. Defect states have a key role in the different steps of the photoelectrochemical process. RHE, reversible hydrogen electrode; TOF, turnover frequency; VB, valence band maximum.

---

**Types of point defects**

Points defects can be broadly classed into two types: first, intrinsic species such as vacancies, when an atom is missing from a regular crystal site, or interstitials, when an atom occupies a crystal site that is normally empty; or second, extrinsic species, including the unintentional incorporation of impurities or deliberate doping. As discussed below, trapped electrons in the form of polarons can also be considered a distinct type of electronic defect.

**Concentration of point defects**

A thermodynamic analysis can yield insight into the concentrations of particular defect species that we might expect in a crystal. The configurational entropy (ΔS) of a static perfect crystal is zero, as each atom is located in its ideal crystallographic site. The consequence is that at non-zero temperatures, the free energy of a crystal (ΔG = ΔH – TΔS, where ΔH is the enthalpy and T is temperature) can be decreased by introducing configurational entropy in the form of defects. A balance emerges between the enthalpic cost of breaking chemical bonds and the entropic gain of forming more imperfections. The resulting exponential dependence means...
that the equilibrium concentration (c) of a defect species is highly sensitive to its energy of formation (ΔHf):

\[ c = N \exp \left( \frac{-\Delta H_f}{k_B T} \right) \]  

(1)

where N represents the density of associated atomic sites in the crystal and \( k_B \) is the Boltzmann constant.

What makes the situation more complex is that the defect formation energy is not a constant, but is a function of the atomic and electronic chemical potentials. In this way, the concentrations of different defect species are coupled and the defect profile of a particular sample will depend on its history, including the growth environment (such as the oxygen partial pressure) and thermal processing (such as the annealing time or quenching rate)\(^\text{63,64}\). Let us take the example of the oxygen vacancy — a ubiquitous defect species in metal oxides. Vacancies can be formed by exchange of oxygen with the atmosphere, as described by \( \text{O}_2 + \frac{1}{2} \text{O}_2 \rightarrow \text{V}_O \). Each oxygen vacancy (\( \text{V}_O \)) has the potential to be a double donor (\( \text{V}_O^{2+} \)) that will transfer two electrons to the conduction band (n-type donor)\(^\text{65-67}\). The formation energy and equilibrium concentration of this defect will be sensitive to the annealing temperature, the partial pressure of oxygen, as well as the presence of other electron-donating defects in the host crystal. Raising the partial pressure during high-temperature annealing and the presence of an electron-rich environment (n-type) will increase the formation energy and thus decrease the concentration of oxygen vacancies in the bulk. The semiconductor α-Fe₂O₃ has a concentration of oxygen sites in the crystal of ~6 × 10²¹ cm⁻³. Assuming a typical defect formation energy of 1 eV, an annealing temperature of 873 K will yield a moderate net vacancy concentration of 6 × 10¹⁷ cm⁻³.

**Energy levels in the bandgap**

From an electronic viewpoint, the biggest consequence of defect formation is a change in the electroneutrality conditions of the metal oxide. An ideal stoichiometric material has a filled valence band and an empty conduction band, which are separated by a bandgap sufficiently large to suppress thermal excitations of electron and hole charge carriers. The incorporation of charged defects is associated with the creation of new energy levels in the bandgap. The position of sub-bandgap defect levels is related to the energy required to exchange charge.

---

**Fig. 1 | Point defects in a metal oxide.** a) Crystal field splitting of d electrons in an octahedral (\( \text{O}_6 \)) and square pyramidal (\( \text{C}_3^+ \)) coordination environment. b) Molecular orbital (MO) diagram describing the building blocks of a transition metal oxide. c) Calculated electronic density of states (DOS) of WO₃. d) Types of point defects that form in metal oxides. e) Oxygen vacancy in an oxide for which the loss of a dangling bond upon ionization results in a structural relaxation. f) Electronic band diagram of an oxygen vacancy in a metal oxide. The defect level is close in energy to that of the metal atomic orbitals (AOs) and can be shifted by structural relaxations, as well as defect–defect interactions. g) Charge density of a deep vacancy state (left) and the corresponding band structure (right). h) Charge density of a shallow vacancy state (left) and corresponding band structure (right). Relaxation of the local structure around the vacancy can induce a transition from a localized (deep) to a delocalized (shallow) state. CB, conduction band; VB, valence band. Panels g and h adapted with permission from Ref. 75, Royal Society of Chemistry.
carriers (electrons or holes) with the valence and conduction bands. For example, a singly charged oxygen vacancy ($V_{O}^{+}$) can be further ionized to inject an electron into the conduction band:

$$V_{O}^{+} \rightarrow V_{O}^{2+} + e^-$$

(2)

or it can arrive in the same final charge state by capturing a hole from the valence band:

$$V_{O}^{+} + h^+ \rightarrow V_{O}^{*}$$

(3)

The energy required for each ionization ($\Delta E$) corresponds to the separation between the defect level and the valence or conduction band edge of the crystal. The location of ionization levels in a semiconductor bandgap is of major interest for the optimization of energy conversion systems because it determines the beneficial or detrimental role of the defect. We distinguish two cases: shallow and deep states. A shallow state describes the case when a defect level lies close (usually within $k_B T = 26 \text{ meV}$ at room temperature) to the band edge and can be thermally ionized. A shallow state is often associated with a defect wavefunction that is delocalized over tens to thousands of unit cells. In the hydrogenic limit, for a charged defect in a dielectric host, the shallow state energy is determined by the effective mass ($m^*$) and the dielectric constant ($\varepsilon$) of the host (in atomic units):

$$\Delta E = \frac{m^*}{2\varepsilon^2}$$

(4)

$\Delta E$ values of 1–50 meV are common in tetrahedral semiconductors such as Si and ZnO. As a consequence, shallow states typically support doping by increasing the concentrations of free carriers in the bands and, thus, the electrical conductivity.

By contrast, a deep state refers to the case when the ionization energy is larger than the thermal energy and is often associated with spatially localized defect wavefunctions on one or a small group of atoms. Deep states are usually formed following a large chemical perturbation, such as a physical vacancy in the structure or an elemental substitution with a large size or orbital mismatch (for example, replacing O by As in a metal oxide). Deep states are often associated with large structural distortions. In optoelectronic and photocatalytic devices, deep states can act as centres for trapping or the recombination of charge carriers and can also pin the Fermi level to a limited energy range inside the bandgap, as discussed below.

We can make a further distinction concerning the defect levels. Optical levels correspond to a rapid vertical excitation (that is, subject to the Franck–Condon principle), during which there is no change in the final defect geometry. By contrast, thermal levels correspond to a slower process, during which the defect geometry has time to transform, as probed using capacitance techniques such as deep-level transient spectroscopy. Therefore, optical levels are deeper than the corresponding thermal levels by an amount that corresponds to the structural relaxation energy of the particular defect species.

From a chemical viewpoint, the depth of the defect states in the bandgap can be understood in terms of the frontier orbitals of the host crystal and the coordination environment of the defect site. The formation of, for example, an oxygen vacancy will create a void with dangling bonds on the adjacent metals that, to a first approximation, will have an energy similar to that of the metal atomic orbitals (FIG. 1e,f). Consequently, the formation of shallow defect states will be favoured when the associated orbitals are located energetically close to the edge of the bands or even within the bands. In addition, weak hybridization between the dangling metals, owing to a large metal–metal distance that reduces orbital overlap and a low coordination environment at the defect site, will favour shallow defect formation.

FIGURE 11 exemplifies the formation of new levels in the bandgap owing to vacancy formation. For example, in WO$_3$, the generation of an oxygen vacancy results in two dangling W bonds with an additional electron each (that is, reduction of W(vi) to W(v)). This configuration leads to the formation of a doubly occupied energy level deep in the bandgap. However, the relaxation of bond lengths around the vacancy to minimize Coulomb repulsion increases the W–W distance and pushes the defect level closer to the conduction band and facilitates the injection of carriers. FIGURE 1g shows a case in which charge is localized around the vacancy and a deep level is formed. In the absence of a strong structural distortion, the charge delocalizes across neighbouring sites, concomitantly forming a shallow defect level (FIG. 1i). When the continuum of states and the wavefunction of the defect overlap, the defect is considered to be resonant with the band.

Polarons

An important consequence of point defect formation is the introduction of excess charge (electrons or holes) into a crystal. For example, doping α-Fe$_2$O$_3$ with Ti or Sn increases the concentration of electrons in the host material and enhances its n-type character. Although defects can induce structural relaxations to accommodate the introduction of the dopant atoms, the excess electronic charge can also induce its own lattice distortion. In particular, if the solid displays a strong dielectric response (that is, it is highly polarizable), the excess charges (electron or holes) can lower their energy by displacing the surrounding atoms through electron–phonon interactions, forming a quasi-particle known as a polaron. As with other point defects, a polaron introduces an energy level in the bandgap (FIG. 2a,b). If the charge and its structural distortion are strongly bound to another defect (such as a vacancy), the polaron can also modify this defect level. For example, electron polarons in BiVO$_4$, that are localized on V centres change the effective oxidation state from V(v) to V(iv) and increase the V–O bond length by ~0.1 Å. Attaining such a distortion might require an activation energy (FIG. 2b). In addition, different configurations might exist for polarons near a defect (vacancy) (FIG. 2c), which complicates the characterization of the localized state both in the bulk and at the surface.

From an energetic viewpoint, the driving force behind polaron formation is a balance between the
potential energy gain upon distorting the structure and the kinetic energy loss associated with carrier localization. Minimization of the total energy of the additional charge in a dielectric continuum yields the polaron binding energy:

\[ E_p = \frac{1}{4\pi^2} \frac{m^*e^4}{2\hbar^2 \varepsilon_{\text{eff}}^2} \]  

(5)

where \( e \) is the electronic charge and \( \varepsilon_{\text{eff}} \) is the effective dielectric constant obtained from the static \( (\varepsilon_s) \) and high-frequency \( (\varepsilon_{\infty}) \) dielectric constants:

\[ \frac{1}{\varepsilon_{\text{eff}}} = \frac{1}{\varepsilon_s} - \frac{1}{\varepsilon_{\infty}} \]  

(6)

The polaron binding energy refers to the energetic stabilization of a polaronic state with respect to a bare valence-band hole or conduction-band electron. Microscopically, polaron formation is dictated by the coupling between charge carriers and the vibrational degrees of freedom. Within the Landau–Pekar model, which considers only the characteristic frequency of a longitudinal optical phonon \( (\omega_o) \), the strength of this coupling, \( \alpha_p \), is

\[ \alpha_p = \frac{e^2}{\hbar^2} \left( \frac{1}{\varepsilon_{\infty}} - \frac{1}{\varepsilon_s} \right) \sqrt{\frac{m^*}{2\omega_o}} \]  

(7)

Depending on the strength of the electron–phonon coupling, and the extent of the structural distortion, polarons are categorized as large or small. Large polarons (or Fröhlich polarons) exhibit weak coupling and the distortion extends over more than a single unit cell, whereas small polarons (Holstein polarons) are characterized by
strong coupling and distortions that extend over less than a single unit cell with wavefunctions localized to dimensions of the order of interatomic distances.

Polaron formation is favoured in systems with large dielectric constants and heavy effective carrier masses (equations (5) and (7)). Evaluation of the electronic structure can therefore provide an indication of the tendency of a material to form polarons. In general, poor overlap between atomic orbitals will result in weakly dispersive bands and large effective masses for carriers that will favour localization. Metal oxides often fall into this regime, as they have frontier orbitals that are made up of weakly interacting orbitals and a high density of states around the band edges.

Such an electronic structure favourable to polaron formation often occurs in transition metal oxides. Indeed, one of the most studied systems in photocatalysis is TiO₂, which is composed of Ti(IV) 3d⁰ ions and in which electron paramagnetic resonance studies have demonstrated the existence of localized electron states in the form of Ti(III) 3d¹ ions. Interestingly, in TiO₂, polaron formation depends on the polymorph; in particular, rutile TiO₂ is known to easily form small electron polarons, whereas increasing evidence suggests that anatase TiO₂ sustains large electron polarons characterized by shallow states below the conduction band. Such a difference in polaron formation can explain the distinct behaviours displayed by anatase and rutile in optoelectronic applications. Similarly, polarons have also been reported to affect other oxides used as photocatalysts, electrocatalysts or co-catalysts, including WO₃ reported to affect other oxides used as photocatalysts, electronic applications. Similarly, polarons have also been reported to affect other oxides used as photocatalysts, electrocatalysts or co-catalysts, including WO₃ reported to affect other oxides used as photocatalysts, electronic applications. Similarly, polarons have also been reported to affect other oxides used as photocatalysts, electrocatalysts or co-catalysts, including WO₃ reported to affect other oxides used as photocatalysts, electronic applications.
**a Solid–liquid phase equilibration**

Charge transfer

Ionizable defects

**b Light harvesting**

Absorbance

300 500 700 900 1,100 1,300 Wavelength (nm)

0.5 1.0 1.5 2.0 2.5 Absorbance

LMCT

MMCT

CoII → CoIII

CoII → CoIII

**c Charge transport**

Band (delocalized)

μ = 1,000 cm−1 V−1 s−1

DOS

Defect (localized)

μ = 0.1 cm−1 V−1 s−1

**d Charge localization**

M3d

O2p

O2s

Localized

Delocalized

**e Charge extraction**

TOF = 0.1–1 s−1

Electron transfer

Impact of defects on photocatalysis

Electric field build-up. Semiconductor photocatalysis is enabled by the semiconductor–liquid junction. When the semiconductor is immersed in the electrolyte, charge transfer across the interface takes place until the electrochemical potentials in the solid and the solution are equal, that is, until the Fermi level of the semiconductor (E_F) and the redox potential of the electrolyte (E^°) equilibrate. In the case of an n-type semiconductor with E_F > E^°, electrons are transferred to the liquid. The larger number of energy states available in the liquid phase means that the energy levels shift mostly within the semiconductor, causing the bands to bend. Through the equilibration process, charged defects play an important part by acting as the source of charge. Beyond the surface atomic layer, charges depleted away from the surface towards the bulk form a depletion region or depletion layer of width W_D (Fig. 4a). The most important consequence of the equilibration process is the formation of a junction that can assist in charge separation and prevent electron–hole recombination events following the reaction site.

For the equilibration with the liquid, all steps are common with other solar energy conversion technologies, such as solar cells. However, the charge-extraction step imposes severe and unique requirements on catalytic materials. From a band structure viewpoint, extraction in a solar cell involves the collection of charges at contact electrodes, whereas the equivalent process in an oxide-based water-splitting device is the four-electron oxidation of water. For such selective chemistry to occur, the frontier orbitals must have the right orbital composition and suitable energetics (that is, in the electrochemical scale, the energy of the VBM must be greater than the redox potential of the species in solution, E_{VBM} > E^° (ox/red), for oxidation to occur). From a kinetic viewpoint, charge collection in a solar cell occurs on the nanosecond-to-millisecond timescale, whereas water oxidation has been measured to occur on the timescale of seconds, meaning that the photoelectrochemical system requires carriers that are 10^3–10^9 times longer-lived as well as mechanisms for their accumulation near the reaction site.
illumination. The strength of the interfacial field can be controlled by changing the density of ionized dopants (N_d), and thus through defect engineering, as well as by applying an external potential. The width of the space charge layer is

\[ W_{SCL} = \left( \frac{2e \varepsilon_0}{e N_d} \Delta \vartheta \right)^{1/2} \]  

where \( \varepsilon \) and \( \varepsilon_0 \) are the material and vacuum permittivity, respectively, and \( \Delta \vartheta \) is the difference between a given applied potential and the potential at which the bands are flat and there is no excess charge in the semiconductor (termed the flatband potential). FIGURE 4a shows the correlation between the width of the space charge layer and the density of V⁴⁺ and V⁵⁺ states associated with oxygen vacancies in BiVO₄.

**Light absorption.** A major approach to lowering the bandgap of photocatalytic oxides and thus enhancing absorption in the visible range is to introduce impurities. Doping with metals has been used since the early work on photocatalysis with Cr(iii)-doped and V(iv)-doped TiO₂, as well as in studies of SrTiO₃ doped with Mn, Ru or Rh (REFs 124, 125). Similarly, the replacement of oxygen atoms with non-metals has been extensively explored, and N-doped TiO₂ and N-doped tantalates are prime examples of materials in which the absorption has been enhanced to reach up to 500–600 nm (REFs 126–128). Another common strategy

**Fig. 4 | Role of defects in solid-liquid equilibration, light absorption and charge transport.** a | Semiconductor–liquid junction in n-type BiVO₄. During equilibration with the electrolyte, filled V⁴⁺ states associated with oxygen vacancies (OV; orange circles) are emptied (that is, oxidized; empty circles), leading to the formation of a space charge layer of width (W_SCL) (left). Positive applied potentials (E) increase the concentration of V⁴⁺ states and enlarge the space charge layer (right, where AOD is the difference in optical density, which is calculated here as the optical signal at each potential minus that at the open-circuit potential (E_oc)). b | Ground-state optical absorption of SrTiO₃ (STO) measured in diffuse reflectance mode at different doping densities (left) and the associated band diagram for Rh³⁺:STO (right). The diagram exemplifies the role of defect engineering in controlling the optical properties of solids. The black trace shows the undoped sample. Doping with Rh³⁺ increases the absorption at long wavelengths (purple). The controlled addition of reduced-valence Rh³⁺ modulates the optical density (orange and blue traces). c | Absorption spectra (1-T-R, where R is reflectance and T is transmittance) of defect-poor and defect-rich WO₃. The oxygen vacancies increase the optical density in the visible–near-infrared region; however, light absorption at those wavelengths does not contribute to the water oxidation photocurrent, as shown by the negligible incident-photon-to-current conversion efficiency (IPCE) in this range. d | Absorption coefficient of a 7-nm-thick α-Fe₂O₃ thin film (black) compared with the estimated upper and lower (dashed orange lines) and average (solid orange line) proportion of carriers that are active for water oxidation. e | Configuration coordinate diagram that exemplifies the principle of electron transfer in the small-polaron regime. The reactants Fe²⁺/Fe³⁺ and products Fe²⁺/Fe³⁺ have equilibrium configurations of Q_a and Q_b, respectively. Electron transfer requires a thermally induced event that excites the sample to an activation energy (E_a-hop) and brings the system to an intermediate configuration at the crossing point (Q_c). CB, conduction band; VB valence band. Panel a adapted with permission from REF. 13, American Chemical Society. Panel b adapted with permission from REF. 14, Royal Society of Chemistry. Panel c adapted from REF. 15, CC BY 3.0. Panel d adapted from REF. 16, Springer Nature Limited.
relies on the formation of oxygen vacancies, which has been used to increase the absorption of wide-gap materials, including ZnO (REFS 129,130), SnO2 (REF 131) and WO3 (REF 132), and even to produce black TiO2 (REFS 133,134) and, more recently, black BiVO4 (REF 135).

The mechanism by which low dopant concentrations enhance optical absorption is generally the introduction of sub-bandgap states that enable longer-wavelength optical transitions between defect and band states or between the defects themselves. This effect is shown in FIG 4b for the case of Rh4+-doped SrTiO3, in which the Rh ions substitute the Ti centres 136. The introduction of Rh4+ ions with a 4d4 configuration opens up new d–d and charge-transfer transitions that increase the absorption at wavelengths beyond 400 nm (REF 137). Further doping with La reduces the valence state of Rh4+ through charge compensation, lowering the absorption in the visible region. This sensitivity to the presence of dopants emphasizes that fine control of the intragap density of states is required to modulate the absorption properties on demand. Engineering of the absorption characteristics can be achieved with the principles outlined above. As a first approximation, the defect state will have an energy similar to that of the atomic orbital of origin (for example, the atomic orbital of the cation surrounding an anion vacancy) and the energy level can be tuned by controlling the coordination and bonding of the defect state.

The power of defect control is seen in the case of WO3, which changes colour from transparent yellow to dark blue with an increase in the density of oxygen vacancies — a property that is exploited for photocatalytic 137,138, photochromic and electrochromic applications 139,140. The blue colour has primarily been attributed to light with an increase in the density of oxygen vacancies — a property that is exploited for photocatalytic 137,138, photochromic and electrochromic applications 139,140. The blue colour has primarily been attributed to light absorption of W 5d states associated with an electron localized in the vicinity of the vacancy, and W 5d states caused by an electron and lattice distortion located further away from the vacancy (that is, an electron polaron unbound from the source defect) 141. The blue colour has primarily been attributed to light absorption of W 5d centres trapped next to the vacancy (bound polaron) and to the absorption of unbound polaronic states. Both of these processes involve charge transfer between W ions, such as from W(v) to W(vi). The spread of defect levels, and thus the optical transition available, can be tuned by favouring defect–defect interactions under highly reducing growth conditions. The defect levels for single oxygen vacancies are located up to 0.7 eV below the conduction band, but the formation of defect clusters alters the coordination of the vacancy and its bonding environment. This clustering widens the range of possible charge states and substantially broadens the intragap energy distribution, which can reach as far as 0.9 eV below the conduction band with a concomitant increase in light absorption at long wavelengths 141. Similar defect–defect interactions are known to have an important role in solar cell materials, such as Cu(In,Ga)(S,Se)2 (REF 142), in which they have also been used to control the tolerance to undesired defects. Modulation of defect–defect interactions through control of the concentration and growth environment offers a strategy to tune the density and distribution of intragap states in catalytic oxides.

Although enhancing optical absorption is necessary, it is important to recognize that not all photogenerated charges are ultimately able to drive a desirable photocatalytic reaction. First, charges generated through the direct photoexcitation of defect states, such as oxygen-vacancy absorption bands in highly oxygen-deficient metal oxides, might lack the necessary driving force to perform thermodynamically demanding reactions 143. For example, for WO3, the absorption enhancement at ~750 nm due to defect states does not contribute to the water oxidation photocurrent (FIG 4c). Second, even charges produced by an intrinsic optical transition might not uniformly contribute to catalysis. For example, the average proportion of charges able to drive water oxidation in α-Fe2O3 photoelectrochemical cells is substantially smaller than unity (FIG 4d). Such phenomena have been ascribed to the formation of photoinduced polarons, which, as described below, limit photoconversion efficiencies 144. The generation of inactive charges is also disproportionately large near the absorption onset, which similarly to observations in early studies, has been attributed to a decreased activity of the ligand field transitions that dominate in the visible spectral range 140–144. These observations indicate the need to carefully engineer oxides to ensure that light absorption generates excited states (whether band-like or polaronic) that can thermodynamically drive catalysis. Moreover, the strong wavelength dependence of photocatalysis highlights the need to better understand how absorption couples to surface catalysis, as well as the intermediate steps such as charge transport and localization.

**Charge transport.** One of the main goals for defect engineering of oxide photocatalysts has been to increase the conductivity of majority carriers. For an n-type semiconductor, the electron conductivity (σ) depends on the concentration of electrons (n) and their intrinsic mobilities (μ):

\[ σ = enμ \]  

Thus, the introduction of high-valence dopants that create shallow defects increases the conductivity by increasing the electron concentration in the conduction band. This effect has been extensively explored; for example, α-Fe2O3 has been n-doped with Ti (REFS 149–151), Si (REF 152), Sn (REFS 135,136), and Ge (REF 134), among others, and BiVO4 has been doped with W and Mo, with important implications for their photoelectrochemical performance 150–152. In addition to metals, the most common source of doping is oxygen vacancies, which are often responsible for the much higher intrinsic doping densities of metal oxides compared with elemental semiconductors 153.

The underlying doping mechanism is not always fully understood, in particular when the excess carriers form polarons. In the small-polaron regime, charge transport occurs via hopping between adjacent atomic sites as opposed to band-like transport 154. This charge transport
mechanism leads to considerably lower carrier mobilities (for example, ~0.1 cm²V⁻¹s⁻¹ in BiVO₄) than in traditional semiconductors (for example, ~1,350 cm²V⁻¹s⁻¹ for electrons in Si)⁹⁰,⁹¹. The reason for such slow transport is that movement of a polaronic charge from site to site requires a structural rearrangement with an activation barrier \( E_{\text{act}} \) with the polaron mobility given by \( \mu \propto e^{-E_{\text{act}}/k_B T} \) (REFS.¹⁶⁰,¹⁶¹). For electron transfer to occur, the reactants must first be excited (for example, thermally) to the activation barrier and assume an intermediate configuration (FIG. 4e). We note that in the diabatic regime, the activation barrier for hopping is related to the polaron binding energy, \( E_{\text{p}} = 2E_{\text{hop}} \) (REFS.¹⁶⁰,¹⁶¹) and, thus, the more strongly bound the polaron, the lower its mobility.

BiVO₄ is an example of a photocatalyst in which electron transport is considered a limiting factor. Although oxygen vacancies are thought to be the source of n-doping, several spectroscopic studies have found that vacancies give rise to donor states as much as 0.6 eV below the conduction band. Such deep character should prevent these states from being ionized, and, thus, should render them ineffective n-dopants. However, recent theoretical work has explained this discrepancy, suggesting that although the oxygen vacancy states are indeed deep, their ionization to more conductive polaronic states is only of the order of 0.1–0.3 eV (REF.¹⁶³). This vacancy-to-polaron ionization value is consistent with experimentally measured activation energies for carrier de-trapping.¹⁶⁵ Polaron transport is also affected by the differences in the concentration of oxygen vacancies between the bulk and surface, as observed using transmission electron microscopy, as well as by variations in the adopted geometrical distortions and configurations, all of which can notably alter the mobility and energetics of the respective polarons.¹⁶⁶

In addition to the effects of vacancies, the presence of other imperfections in the crystal can also change polaron transport. For example, doping of α-Fe₂O₃ with Si has been reported to increase the conductivity by increasing carrier concentrations and also by increasing the mobility of small polarons.¹⁶⁷ The incorporation of Si enhanced polaron transport by changing bond lengths, altering hopping probabilities in the system and lowering the activation energy for hopping (FIG. 4e). Crucially, these positive effects of Si are not general for all dopants and might depend on the radius of the foreign atom and the extent to which it distorts the crystal. For example, Sn-doping had a negative effect on polaron mobilities in α-Fe₂O₃, increasing the activation barrier for hopping.¹⁶² However, Sn-doping is used in active α-Fe₂O₃ photocatalysts because the detrimental effect on polaron transport can be compensated by an increase in electron concentration. These results show that doping is a more complex process than the simple injection of carriers into a band. The countering effects of doping with atoms such as Si and Sn might prove useful in optimizing the carrier density and polaron transport through careful co-doping strategies.

Pump–push photocurrent experiments of α-Fe₂O₃ photoelectrochemical cells have also demonstrated that the formation of electron polarons within less than a picosecond affects charge extraction in the photoelectrochemical device, because there is a large population of polaronic charges that recombine before being extracted. Similarly, ultrafast terahertz conductivity measurements in BiVO₄ have reported a combined electron–hole peak mobility of 0.4 cm²V⁻¹s⁻¹ at ~1 ps that decreases by ~70% within the first few hundreds of picoseconds. This loss signifies a substantial loss of free or relatively unbound carriers, which can result from carrier localization. Time-resolved microwave conductivity measurements on the nanosecond-to-millisecond timescale have yielded similar low mobilities for undoped BiVO₄, further supporting the localized picture. However, the observed carrier lifetimes of ~40 ns lead to a relatively long calculated carrier diffusion length of ~70 nm and suggest that despite the low mobility, photogenerated charge carriers can access interfaces for reactions in this material.

Impact of charge localization on defects. Even when a beneficial dopant is identified, low doping efficiency is a problem that can affect many oxides. This issue happens when only a small fraction of the dopants is active (ionized). Low doping efficiency can occur either from dopant levels being too deep in the bandgap or the migration of dopants to crystal boundaries, where their behaviour changes and they become inactive. Deep defect states act as trapping sites for photogenerated charge carriers. Such a trapping process has two main consequences: first, it can immobilize one type of charge carrier through selective trapping and, second, it enables non-radiative electron–hole recombination through a Shockley–Read–Hall mechanism. The first process can be beneficial, because trapped carriers can exhibit very long lifetimes, as observed in different polymorphs of TiO₂ as well as in α-Fe₂O₃ and WO₃. In the absence of other pathways to increase carrier lifetimes, selective trapping can aid charge separation and ultimately improve catalysis.

The complexity of the localization process is exemplified for WO₃ (FIG. 5a). In WO₃, a large part of the W 5d vacancy state distribution is energetically too far from the conduction-band edge for thermal excitation. As a consequence, intragap states in WO₃ are largely occupied with electrons (in the form of reduced W(V) centres) and, thus, can act as trapping sites for minority carriers (photogenerated holes). Such trapping enables a relaxation pathway for the hole, through which the energy of the hole is lowered by oxidizing a W(V) centre, losing most of its oxidative driving force in the process. Importantly, hole trapping does not immediately lead to annihilation of the counterpart photogenerated electron. This behaviour is corroborated by measurements on longer timescales, which show a substantial gain in lifetime of up to several milliseconds. Hence, the localization process provides a mechanism by which carrier lifetimes can be extended at the expense of carrier energy or driving force.

The localization process can be tuned by controlling the distribution of intragap defect states and, as expected, highly oxygen-deficient WO₃ with a broad defect distribution exhibits faster trapping than the
near-stoichiometric analogue. This behaviour is exemplified in the transient dynamics of WO$_3$ (Fig. 5b), in which the appearance of the negative signal is associated with hole trapping$^{74,171}$. The same behaviour has been observed in BiVO$_4$: within $<200$ fs of photoexcitation, a fraction of the photogenerated holes relax into intragap V(IV) states associated with oxygen vacancies, oxidizing them into V(V) states. Relaxation into such states, which are located at approximately 0.6–0.8 eV relative to the CBM and are thus more negative than the thermodynamic water oxidation potential of 1.23 V versus the reversible hydrogen electrode, renders these photogenerated holes inactive for water oxidation$^{163}$. It is important to note that although a decrease in driving force associated with carrier relaxation means that the charges might become unreactive towards highly demanding, multi-electron reactions, the concomitant increase in lifetime might be advantageous to drive less energetically demanding processes, such as pollutant degradation. Hence, engineering of defect distributions and defect concentrations might open up a route towards tuning oxides to specific catalytic reactions. In this sense, thermodynamically facile reactions might be more tolerant of and even benefit from high defect concentrations, whereas more demanding processes might require lower defect densities. Indeed, the performance
optimum for water oxidation in WO$_3$ is at only ~2% oxygen vacancies$^{41}$. 

**Impact of charge localization on polarons.** The formation of polarons upon photogeneration provides another pathway for carrier energy minimization. In BiVO$_4$, hole polarons are reported to be weakly localized and located only 0.1 eV above the VBM; consequently, polaronic holes can drive water oxidation. By contrast, electron polarons are strongly localized on V atoms and exhibit stabilization energies as large as 0.9 eV relative to the CBM$^{44}$ (fig. 5c). The electron-polaron level is more positive than the $\text{H}^+\text{H}_2$ redox potential, rendering polaronic electrons unreactive towards proton reduction and limiting the ability of BiVO$_4$ to drive the overall water-splitting reaction. Moreover, the electron-polaron level is positioned close to the $\text{H}_2\text{O}/\text{O}_2$ level, indicating that recombination of electron–hole polarons can effectively compete with charge transfer to the electrolyte$^{44}$. We note that polaron binding energies depend on temperature, the presence of other defects and the physical location within the oxide (that is, surface versus bulk). Consequently, there is a spread of absolute values reported in the literature.

In α-Fe$_2$O$_3$, the stabilization energy for electron polarons is also large (~0.5 eV)$^{48}$ and decreases the optical gap (2.2 eV) to an effective electronic gap of only ~1.75 eV (REF. 173). X-ray photoemission studies have suggested that such a large decrease in the energy of the electronic state limits the attainable photovoltage in α-Fe$_2$O$_3$ devices and thus restricts its ability to drive photochemical reactions$^{174}$. Similar to what has been observed in other systems such as organic semiconductors, it might be possible to decrease such losses through polaron engineering. One route could be to decrease the binding energy or to increase the activation barrier for polaron formation through crystal engineering (fig. 2b). This approach would decrease the probability of polaron formation and minimize the decrease in the energy of the electronic state, easing the penalty on the photovoltage. Moreover, a lower polaron binding energy would also induce a lower activation barrier for carrier hopping, increasing transport. Guiding the redesign of polaronic extended solids is non-trivial, but classical concepts derived from coordination chemistry might be of use. Indeed, experiments have shown that photoinduced polaron formation depends on molecular-scale parameters, such as ligand field strength and the metal coordination environment$^{178}$. In this context, strategies such as the use of metal oxohydroxides or multinary oxide systems such as ferrites, which are being explored for photocatalysis$^{175}$, could help control polaron formation and its impact on catalysis.

The band alignment of several frequently used oxides and their respective intragap states are shown in fig. 5d. Although it has been established that defects affect the performance of these oxides, it remains an open question whether the charge–carrier localization process can be overcome to increase catalytic yields or whether they are unavoidable, and potentially even necessary, to achieve sufficient lifetime gains. Interestingly, the trade-off between lifetime gain and energy loss observed in oxides is not unique to solids but is a central feature in the operation of the natural photosynthetic system. In photosystem II (fig. 5d), a series of redox cofactors enable the spatial separation of the chlorophyll excited state$^{174}$. The result is a lifetime gain of 10$^6$ at the expense of an energy loss of 700 meV, equivalent to 40% of the energy of a red photon$^{174}$. Such cofactor states can be approximated as the ‘molecular analogue’ of intragap states in extended solids. However, the photosystem can modulate the redox cofactors to regulate the backwards and forwards electron transfer and thus displays considerable adaptability and photoconversion flexibility$^{176}$. To achieve high photocatalytic efficiencies, the challenge is to outperform nature by decreasing or better controlling the energy losses in artificial systems. Learning to minimize unwanted defects and how to energetically and spatially control defect populations could enable the necessary fine-tuning and even reveal new materials functionality$^{177,178}$. Achieving this goal might require the exploitation of defect–defect and polaron–defect interactions to tune energy distributions and the tolerance to unwanted defects.

**Surface chemistry.** The last stage in the photocatalytic process is the extraction of charges from the active material. This final step marks the start of the catalytic part of the mechanism in which photogenerated charges are extracted through a chemical reaction that occurs at the solid–liquid interface. Many reactions of interest, such as proton reduction or water oxidation, are multi-electron reactions and involve the formation and accumulation of several reactive intermediates, which means that the completion of the catalytic cycle and final product formation are slow (milliseconds to seconds)$^{179,180}$. The role that defects have at this stage is challenging to establish. Experimentally, it is difficult to characterize the surface under operation to elucidate, at an atomic level, how the surface changes upon contact with the electrolyte or under different working conditions. This problem is not a minor issue; for example, in oxide photoanodes the mechanism of the oxygen evolution reaction (OER) has been proposed to change as a function of pivotal parameters such as the illumination intensity$^{181,182}$. From a simulation viewpoint, it is difficult to build realistic models that describe the influence of solvation and temperature fluctuations on the frontier orbitals owing to the high computational cost. Such information is essential to explain reactivity and defect formation at the interface. Consequently, the surface of photocatalysts under working conditions is, to a great extent, uncharted territory. Nonetheless, new experimental methods and computational electrochemistry approaches are increasingly providing insight into the solid–liquid interface$^{183–189}$.

One of the biggest consequences of the lack of operational surface characterization is the uncertainty in the energy levels of defects, polarons and even band edges in contact with the liquid. This uncertainty means that, in many cases, it is unclear whether a given state can thermodynamically participate in redox chemistry. Thus, there have been contrasting interpretations of the role of defects, with some reports suggesting that defects have positive effects and mediate catalytic reactions,
whereas others propose that defects are unreactive, detrimental and should be eliminated or passivated\textsuperscript{191–194}. From a reactivity standpoint, defects have several fundamental roles in the surface chemistry. First, defects can form as part of the catalytic cycle and the interaction with the electrolyte\textsuperscript{195,196} (Fig. 6a). Second, the presence of defects can ‘physically’ change the nature of the surface by, for example, exposing new reactive sites\textsuperscript{197} (Fig. 6b). Third, defects can modulate the availability of electronic charge at the surface and the charge transport between surface sites\textsuperscript{195,197,202} (Fig. 6c). Finally, the presence of defects can modulate the energetics of the surface by, for example, altering the energy of reaction intermediates, facilitating the adsorption of molecules and even modifying the magnetic properties of the surface\textsuperscript{198–202} (Fig. 6d).

Given the abundance of oxygen atoms in oxides, the formation of oxygen vacancies is particularly important, especially in water-based electrolytes. In perovskite electrocatalysts, $^{18}O$ isotope labelling mass spectrometry has shown that the $O_2$ evolved during the OER can come from the lattice as opposed to exclusively from water\textsuperscript{203} (Fig. 6e). The activation of lattice oxygen and vacancy formation during the catalytic cycle (with their subsequent ‘re-filling’) has not been extensively studied in photocatalysis, but it is a possible route towards increasing the redox activity. As expected from the discussion above, whether vacancies are generated will depend on the metal–oxygen bonding and the nature of the frontier orbitals. Such effects have been systematically studied in oxide electrocatalysts, revealing that higher bond covalency increases the $O_2p$ character near the Fermi level and favours the participation of lattice oxygen atoms during catalysis\textsuperscript{203,204}. Control over the covalency and an increase in the overlap between the metal and oxygen orbitals can be achieved by lowering the energy of the cation (by using a more electronegative metal) or through incorporation of $4d$ cations with more spatially extended orbitals\textsuperscript{205} (Fig. 1c). The crucial aspect is that in these covalent systems, the band structure allows for the creation of holes in the oxygen band, thus enabling vacancy formation. This feature might also be the case in oxide photocatalysts. For example, photoelectron spectroscopy studies on BiVO$_4$ have shown that upon exposure to water, the V and Bi atoms are reduced and the oxygen content at the surface decreases owing to the formation of oxygen vacancies\textsuperscript{196}. Such findings emphasize the importance of oxygen vacancies at the solid–liquid interface and emphasize the need for surface characterization and labelling studies to expose the role of lattice oxygen in photocatalytic reactions\textsuperscript{206–209}.

In addition to their participation in the catalytic cycle, oxygen vacancies in BiVO$_4$ can alter the surface reactivity towards the OER. This effect has been proposed to occur by two mechanisms, namely by activating the V atoms, and thus increasing the number of active sites, and by changing the energy of adsorbates and increasing hole transfer\textsuperscript{206}. Moreover, the presence of vacancies also increases excess charge at the surface, which has possible implications for the reactivity. In WO$_3$, this surface defect has been calculated to be favourably positioned to transfer electrons to OH, possibly forming OH$^\cdot$. Subsequently, delocalized surface holes might transform OH$^\cdot$ into a reactive OH$^\cdot$ intermediate. Crucially, in WO$_3$, the excess charge due to the vacancy has been proposed to form a delocalized large polaron (1 nm in diameter)\textsuperscript{206} (Fig. 6f). Control over polaron formation at the surface is imperative, as, for example, recent measurements have shown that the rate-determining step in the OER on $\alpha$-Fe$_2$O$_3$ involves a triply oxidized cluster\textsuperscript{210}. The accumulation of three oxidation equivalents on one specific site requires the lateral diffusion of holes over the surface and, consequently, necessitates an efficient transport mechanism (note that strong polaron localization with high binding energies also leads to a higher activation barrier for charge hopping, as discussed above).

The changes that defects such as dopants or vacancies have on the surface chemistry can be more intricate than the addition of excess charge or the exposure of new sites. For example, recent experimental work has shown that tuning the concentration of Ti vacancies in
TiO₂ provides an approach to alter the spin polarization in the system and enhance photocatalytic hydrogen evolution and phenol degradation. An increase in the concentration of Ti vacancies increases the population of electrons in minority spin states. This electronic polarization favours the formation of spin-polarized carriers upon photoexcitation and was suggested to influence the recombination of hydroxyl radical intermediates through spin-restriction processes (that is, a spin-selective electron exchange with OH⁻ species in solution). Such magnetic effects are not exclusive to TiO₂; the use of magnetic fields has been proposed to enhance the OER with Ni–Fe oxide electrocatalysts by favouring the spin-parallel alignment of oxygen radicals during O–O bond formation. These advances suggest that tuning the spin polarization of the active layer, through magnetic fields or defect formation, could provide a new avenue to boost photocatalytic reactions.

**Comparison with other photocatalysts**

Although this Review focuses on metal oxides, an awareness of defect states is important for any photocatalyst. We therefore end with a brief comparison to defects in other photocatalytic materials to put the defect chemistry of oxides into perspective.

**Organic photocatalysts**

Organic polymers have become popular photocatalysts because of their high synthetic tunability and are emerging as promising photocatalyst candidates. This high synthetic tunability has sparked the creation of diverse polymeric materials with varying degrees of light absorption, conjugation, porosity and hydrophilicity, among other properties. The most prominent examples of polymeric photocatalysts are carbon nitrides, CNₓHᵧ, which are based on triazine or heptazine units and possess bandgaps of 2.7 eV (ref. 215). Like metal oxides, carbon nitrides have been explored for several catalytic reactions, such as H₂ evolution, O₂ evolution, CO₂ reduction and pollutant degradation.

Although the nature of the frontier orbital is clearly different to that of oxides, increasing evidence suggests that the activity of newly developed carbon nitride photocatalysts is also strongly affected by their defect chemistry in a similar way to oxides. Spectroscopic studies indicate that charge trapping in urea-derived materials results in energy losses of ~1.5 eV (about half the energy of absorbed photons), substantially decreasing the reactivity of photogenerated charges and suggesting that deeply trapped long-lived electrons are unreactive for photoreduction (ref. 216) (FIG. 5c). By contrast, in melamine-derived carbon nitride, long-lived trapped electrons retained sufficient chemical potential to drive proton reduction, even after 12 h in the dark.

Another similarity of organic semiconductors with metal oxides is the tendency to form polarons upon photoexcitation. Polaron formation has been observed in organic photocatalysts and especially in photovoltaics. In general, polarons are detrimental to the dissociation of charge-transfer states and consequently decrease the voltage output of devices. This behaviour parallels that proposed for metal oxide photoelectrochemical cells.

What is most interesting is that, experimentally, efficient organic systems do not require band states for charge separation, indicating that the detrimental influence of polaronic effects can be lessened by materials engineering. Similar synthetic strategies to those developed for organics could potentially be applied to oxide photocatalysts in order to circumvent polaron-associated losses.

Despite the big differences between carbon-based and inorganic photocatalysis, it is apparent that they share similar molecular-scale principles in relation to defect chemistry. These similarities offer an exciting opportunity to draw comparisons between these systems and, for example, to use the synthetic tunability of organics to devise defect-control strategies that are also applicable to oxides.

**Metal halide perovskites**

Organic–inorganic (for example, CH₃NH₃PbI₃) and inorganic (for example, CsPbI₃) halide perovskites are increasingly being explored for photocatalytic applications. In comparison to metal oxides, these systems exhibit higher conduction-band (E_CB < E°(H₂/H⁺)) and valence-band energies (E VB > E°(O₂/H₂O)) and, consequently, are explored for reduction reactions or the thermodynamically facile oxidation of organic compounds.

Generally, metal halide perovskites exhibit relatively low defect formation energies and thus sustain high defect concentrations. However, they are heavily compensated systems in which the high concentrations of positively and negatively charged defects cancel, resulting in low electronic carrier concentrations. Metal halide perovskites have also proved difficult to dope, because these soft crystals can self-regulate their charge by easily forming compensating defects. Another difference to oxides is that such defects tend to be shallow in nature and, thus, relatively benign in the halides. The shallowness of the defects is a consequence of the electronic structure, which greatly differs from that of oxides, in addition to the low carrier effective masses and the large dielectric constants (due to their polarizable structure).

The ‘defect-tolerant’ electronic structure of halide perovskites means that photoexcited charges retain most of their electronic energy — an advantageous feature for photocatalysis. Although there is still no consensus on the exact mechanism, increasing evidence indicates that (large) polaron formation assists in screening the excited states and helps to decrease carrier-recombination rates. The formation of large polarons, as opposed to small polarons with the large binding energies typical of oxides, is due to the stronger orbital overlap that results in lower carrier effective masses. Most importantly, polaronic effects can be engineered, as the softness of the material can be modulated by varying the chemical composition. Metal halide perovskites are examples of materials in which the trade-off between energy loss and lifetime gain, characteristic of photocatalytic oxides and the natural photosynthetic system, can be more readily optimized.

The major limiting factor of metal halide perovskites for photocatalysis is their poor stability, especially in...
polars and salt solutions under illumination. For example, when exposed to water the material tends to suffer surface reconstructions, forming grain boundaries. Mixed halide perovskites tend to undergo phase separation under constant illumination, and the facile formation of iodine vacancies can promote degradation because the excess charge associated with the vacancy reacts with O₂ to form superoxide (O₂⁻). However, further studies are required to unveil the role of defects in photocatalytic perovskites, as this might differ from their role in solar cells. For example, recent work studying the degradation of an organic compound has shown that iodine vacancies are the central species that dictate photocatalytic activity.

**Outlook**

A big part of the challenge facing photocatalysis involves balancing the supply and demand of reactive charges by coordinating the charge-generation and charge-separation steps with efficient charge extraction. Kinetically, this requires the generation of charge carriers that survive long enough to supply a catalytic active site and support high turnover numbers. Energetically, such long-lived carriers must retain sufficient free energy to drive the desired chemical transformation. In a solar cell, this synchronization is relatively simple, and in photosystem II, it is achieved through a series of complex and dynamic structures within the protein framework. Although such structures are difficult to replicate synthetically, defect engineering might provide the chemical complexity and tunability needed to improve artificial systems.

From a materials viewpoint, optimizing the trade-off between lifetime gain and energy loss will require several advances. It will be necessary to go beyond present doping strategies and learn to control sub-bandgap states using defect−defect and defect−polaron interactions. Moreover, polaron-engineering strategies in oxides should be explored by synthetically tuning the metal coordination environment and altering the structural softness. Experimentally, such advances will require the in situ characterization of defects and polarons during operation and an investigation of how the complicated energy landscape that defects generate is affected by the catalytic conditions. Such studies should be guided by detailed theoretical work that pushes beyond the current standard of studying isolated defects in a perfect host at 0 K, instead exploring systems at room temperature and equilibrated with liquid electrolytes, to account for the multiple configurations that vacancies and other defect species can assume, both in the bulk and at the crystal boundaries.

Published online 4 May 2022
61. Brillet, J. et al. Efficient water splitting by nanobrass perovskite tandem cells. Sci. Adv. 2, e1507764 (2016).
62. Chen, C. et al. Highly efficient water splitting by nanobrass perovskite oxide hollow nanospheres as solar water splitting catalysts. Science 343, 990–994 (2014).
63. Qiu, H. et al. Efficient solar-driven water splitting by nanobrass perovskite solar cells. Sci. Adv. 2, e1507764 (2016).
64. Park, J. S., Kim, S., Xie, Z. & Walsh, A. Point defect engineering in BiVO4. J. Mater. Chem. A 2, 12422–12432 (2015).
65. Sai Gautam, G., Senftle, T. P., Alidoust, N. & stoneham, A. M. Small polarons in water splitting: the case of BiVO4. Chem. Commun. 56, 2832–2835 (2020).
66. Biswas, S., Husek, J., Londo, S. & Baker, L. R. Ultrafast electron trapping and defect-mediated recombination in NiO probed by femtosecond extreme ultraviolet reflection–absorption spectroscopy. J. Phys. Chem. Lett. 9, 5047–5054 (2018).
67. Wei, Y., Zhou, Z., Fang, W.-H. & Long, R. Grain boundary facilitates photocatalytic reaction in rutile TiO2, despite fast charge recombination: a time-domain analysis. J. Phys. Chem. Lett. 9, 5886–5889 (2018).
68. Bai, S., Zhang, N., Cao, C. & Xiong, T. Defect engineering in perovskites for solar energy applications. Nano Energy 53, 1–9 (2018).
69. Krüger, F. A. The Chemistry of Imperfect Crystals: Preparations, Structures, Growth and Phase Theory Vol. 1–3 (North-Holland, 1973).
70. de Groot, F. M. F. & Min, J. et al. Oxygen 1s X-ray-absorption edge structures of Zero transition metal oxides. Phys. Rev. B 40, 5715–5718 (1989).
71. Arima, T., Tokura, Y. & Torrance, J. B. Variation of optical gaps in perovskite-type 3d transition-metal oxides. Phys. Rev. B 43, 17006–17009 (1991).
72. Lany, S. Semiconductor transition metal oxides. J. Phys. Condens. Matter 27, 282305 (2015).
73. Huang, J. et al. Efficient light-emitting diodes in catalysts and electrocatalysts. Science 358, 751–756 (2017).
74. Grimaud, A. et al. Double perovskites as a family of highly active catalysts for oxygen evolution in alkaline solution. Nat. Commun. 24, 2549 (2013).
75. Hong, W. T. et al. Toward the rational design of non-precious transition metal oxides for oxygen evolution electrocatalysis. Energy Environ. Sci. 8, 1404–1427 (2015).
76. Calle-Valls, F. et al. Number of outer electrons as descriptor for adsorption processes on transition metals and their oxides. Chem. Sci. 5, 1245 (2015).
77. Nie, X., Wen, S.-H. & Zhang, S.-B. First-principles study of different p-type conductive SrCuO2 and related compounds. Phys. Rev. B 65, 075111 (2002).
78. Greiner, M. T. & Lu, Z.-H. Thin film metal oxides in organic electronics: their electronic structures, work functions and interfaces. npg Asia Mater 5, e55 (2013).
79. Zawadzki, E., Jaccard, K. W. & Rossmeisl, J. Electronic hole localization in rutile and anatase TiO2 — self-interaction correction in DFT. Phys. Rev. Lett. 103, 046402 (2009).
80. Davies, D. W. et al. Descriptors for electron and hole charge carriers in metal oxides. J. Phys. Chem. Lett. 1, 443–448 (2010).
81. Meiggiolaro, D., Ambrosio, F., Mosconi, E., Mahata, A. & De Angelis, F. Polaron in metal halide perovskites Adv. Energy Mater. 11, 2355 (2021).
82. Yang, S., Brant, A. T., Giles, N. C. & Halliburton, L. E. Intrinsic small polarons in rutile TiO2. Phys. Rev. B 87, 155201 (2013).
83. Moser, S. et al. Tunable polaronic conduction in anatase TiO2. Phys. Rev. Lett. 110, 196403 (2013).
84. Moser, S. et al. Unravelling small-polaron transport in the excited state of anatase TiO2 measured by inelastic X-ray scattering spectroscopy. Phys. Rev. Lett. 115, 096404 (2015).
85. Setvin, M. et al. Direct view at excitons in TiO2 rutile and anatase. Phys. Rev. Lett. 113, 086402 (2014).
86. Gerosa, M., Cygi, F., Goboni, M. & Galli, G. The role of defects and excess charges at finite temperatures for optimizing oxygen photoabsorbers. Nat. Mater. 17, 1122–1127 (2018).
87. Kweon, E., Kim, S., Kim, K. & Kim, S. Electron small polarons and their transport in bismuth vanadate: a first principles study. Phys. Chem. Chem. Phys. 18, 12537–12544 (2016).
88. Butler, K. T. et al. Ultrafast carrier dynamics in BiVO4, thin film photoanode material: interplay between free carriers, trapped carriers and low-frequency lattice vibrations. J. Mater. Chem. A, 1851–1865 (2016).
89. Gru, W. et al. Freeing the polarons to facilitate charge transport in perovskites from oxygen vacancies with an oxide 2D precursor. Angew. Chem. Int. Ed. 58, 19087–19095 (2019).
90. Zirnischtz, M. et al. Time-resolved observation of carrier trapping and polaron conductivity in BiVO4. ACS Energy Lett. 1, 888–894 (2016).
91. Mohammadi, A. et al. The electronic structure and the formation of polarons in Mo-doped BiVO4 measured by angle-resolved photoemission spectroscopy. ACS. Adv. Sci. 9, 15660–15614 (2019).
210. Courbon, H., Formenti, M. & Pichat, P. Study of spin polarization of titanium dioxide for efficient photocatalysis. Nat. Commun. 11, 418 (2020).
211. Pan, L. et al. Manipulating spin polarization of pure and doped hematite (0001) surfaces: prediction of Co and Ni as effective dopants for electrocatalysis. J. Am. Chem. Soc. 134, 15396–15399 (2012).
212. Pan, L. et al. Manipulating spin polarization of pure and doped hematite (0001) surfaces: prediction of Co and Ni as effective dopants for electrocatalysis. J. Am. Chem. Soc. 134, 15396–15399 (2012).
213. Jin, H. et al. It’s a trap! On the nature of localised states and charge trapping in lead halide perovskites. Mater. Horiz. 7, 597–610 (2020).
214. Zhu, X.-Y. & Podzorov, V. Charge carriers in hybrid organic–inorganic lead halide perovskites might be protected as large polarons. J. Phys. Chem. Lett. 6, 4758–4761 (2015).
215. Puppin, M. et al. Evidence of large polarons in photothermal band mapping of the perovskite semiconductors CsPbBr3, Phr. Rev. Lett. 124, 206402 (2020).
216. Guzetik, B. et al. Visualization of dynamic polaronic strain fields in hybrid lead halide perovskites. Nat. Mater. 20, 618–623 (2021).
217. Miyata, K. et al. Large polarons in lead halide perovskites. Sci. Adv. 3, e1701217 (2017).
218. Kabakov, I. V. et al. The effect of ionic composition on acoustic phonon speeds in hybrid perovskites from Brillouin spectroscopy and density functional theory. J. Mater. Chem. C 6, 3861–3868 (2018).
219. Limmer, D. T. & Ginsberg, N. S. Photocatalytically produced phase separation in the lead halides is a polaronic effect. J. Chem. Phys. 152, 259091 (2020).
220. Cui, Y. et al. Metal-free activation of H2O2 by g-C3N4 under visible light irradiation for the degradation of organic pollutants. Phys. Chem. Chem. Phys. 14, 1465–1462 (2012).
221. Fabian, D. M. et al. Particle suspension reactors and materials for solar-driven water splitting. Energy Environ. Sci. 8, 2825–2850 (2015).
222. Agter, J. W., Shaver, M. R., Walczak, K. A., Sharp, I. D. & Ardo, S. Experimental demonstrations of spontaneous, solar-driven photocatalytic water splitting. Energy Environ. Sci. 8, 2811–2824 (2015).
223. Jiang, C., Moniz, S. J. A., Wang, A., Zhang, T. & Tang, J. Photocatalytic devices for solar water splitting — materials and challenges. Chem. Soc. Rev. 46, 4465–44660 (2017).
224. Moss, B., Babacan, O., Kafizas, A. & Hankin, A. A review of inorganic photocatalyst development and reactor scale-up challenges for solar hydrogen production. Adv. Energy Mater. 11, 2003286 (2021).